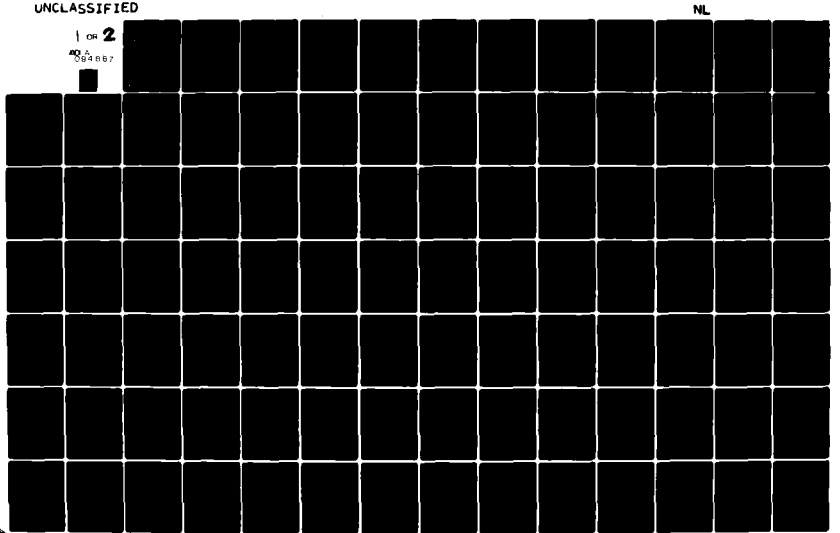
AD-A094 667 OFFICE OF NAVAL RESEARCH ARLINGTON VA F/G 12/1
NAVAL RESEARCH LOGISTICS QUARTERLY. VOLUME 27, NUMBER 4.(U)
DEC 80

UNCLASSIFIED NL

1 OR 2
AD A094867



NAVAL RESEARCH LOGISTICS QUARTERLY

Volume 27, Number 4.

DECEMBER 1980
VOL. 27, NO. 4

OFFICE OF NAVAL RESEARCH

NAVSO P-1278

NAVAL RESEARCH LOGISTICS QUARTERLY

The Naval Research Logistics Quarterly is devoted to the dissemination of scientific information in logistics and will publish research and expository papers, including those in certain areas of mathematics, statistics, and economics, relevant to the over-all effort to improve the efficiency and effectiveness of logistics operations.

Information for Contributors is indicated on inside back cover.

The Naval Research Logistics Quarterly is published by the Office of Naval Research in the months of March, June, September, and December and can be purchased from the Superintendent of Documents, U.S. Government Printing Office, Washington, D.C. 20402. Subscription Price: $11.15 a year in the U.S. and Canada, $13.95 elsewhere. Cost of individual issues may be obtained from the Superintendent of Documents.

Issuance of this periodical approved in accordance with Department of the Navy Publications and Printing Regulations, P-35 (Revised 1-74).

STORAGE PROBLEMS WHEN DEMAND IS "ALL OR NOTHING"*

D. P. Gaver and P. A. Jacobs

Department of Operations Research
Naval Postgraduate School
Monterey, California

ABSTRACT

An inventory of physical goods or storage space (in a communications system buffer, for instance) often experiences "all or nothing" demand: if a demand of random size D can be immediately and entirely filled from stock it is satisfied, but otherwise it vanishes. Probabilistic properties of the resulting inventory level are discussed analytically, both for the single buffer and for multiple buffer problems. Numerical results are presented.

1. INTRODUCTION

The usual storage or inventory problems involve demands imagined to occur randomly, and to be capable of reducing any available stock to zero, or even beyond, when backordering is permitted. Yet in many situations at least one component of total demand is "all or nothing;" that is, it reduces inventory only if it can be entirely satisfied by the inventory present, and otherwise seeks another supplier. Here are examples.

(a) A manufacturer's warehouse is filled with a certain item at the beginning of the selling season; let I denote the initial inventory. Suppose that demands occur as follows: a message is sent requesting that D_1 items be shipped from inventory, but only if the entire order can be filled. That is, the demand is satisfied if $D_1 \leq I$, in which case inventory level is reduced to $I(1) = I - D_1$; while if $D_1 > I$ the inventory remains unchanged and $I(1) = I$. Allowing for no replenishment, the second demand, of size D_2, interacts with inventory $I(1)$, so that it is filled if $D_2 \leq I(1)$, but is not placed if $D_2 > I(1)$. The process continues along these lines until the selling season is over and there are no more demands.

(b) A buffer storage device used to contain messages prior to their batch transmission has capacity I. Messages of length $\{D_i,\ i = 1, 2, \ldots\}$ approach the buffer successively, and are admitted on an "all or nothing" basis, just as was true of demands for physical inventory in (a) above. Once again rejection will occur, and more frequently to large demands (messages) than to short ones.

(c) A system of many buffer storage devices is used to contain messages prior to their batch transmission. Each buffer has capacity I. Messages of length $\{D_i,\ i = 1, 2, \ldots\}$ approach the device and are successively admitted to the first buffer until there is a demand that exceeds its remaining capacity. The first buffer is left forever and the demand that exceeds the first

*This research was supported by the National Science Foundation under NSF ENG 77-09870, ENG 79-01458, and MCS 77-07587, and by the Office of Naval Research under Contract Task NR047-141.

buffer, plus successive demands, applies to the second buffer until one occurs that exceeds the remaining capacity. This demand then applies to the third buffer, and so on. As a result there will be some unused capacity in each buffer. For a similar problem see the paper of Coffman, Hofri, and So [2]. For related, although not identical formulations, see Cohen [3], Gavish and Schweitzer [6], and Hokstad [7].

In Section 2 we will discuss some models for the situations in Examples (a) and (b). We compute such items as the distribution of the amount of inventory left at some time t and the distribution of the times of successive unsatisfied demands.

In Section 3 we next consider a model for Example (c), and derive equations for the limiting distribution of used capacity of a buffer and the expected used capacity of a buffer. It seems to be difficult to obtain simple analytic solutions to these equations, but certain illustrative numerical results are provided.

2. THE ONE-BUFFER INVENTORY PROBLEM

Suppose that demands for available stock occur according to a compound Poisson process: if N_t is the number of demands that occur in $(0, t]$, then $\{N_t; t \geqslant 0\}$ is a stationary Poisson process with rate λ; the sizes of successive demands $\{D_i\}$ are independent with common distribution F. Assume that there are no replenishments of inventory. Let $\{I_t; t \geqslant 0\}$ denote the stochastic process describing available inventory at time t, and let $\{I(n); n = 0, 1, \ldots\}$ be the stochastic process of available inventory following the nth demand. It is apparent from our assumptions that both $\{I_t\}$ and $\{I(n)\}$ are Markov processes.

2.1 Functional Equations for the Amount of Available Inventory

Let

$$\phi(s,t) = E[e^{-sI_t}] \tag{2.1}$$

be the Laplace transform of the available inventory at time t. Similarly, let

$$\psi(s,n) = E[e^{-sI(n)}].$$

Properties of the available inventory can be studied in terms of ϕ and ψ. It may be shown by using conditional expectations that ϕ satisfies the following differential equation.

$$\frac{\partial \phi}{\partial t} = \lambda E\left[e^{-sI_t} \int_0^{I_t} (e^{sx} - 1)\, F(dx)\right]. \tag{2.2}$$

Further, ψ satisfies the following difference equation

$$\psi(s, n+1) = \psi(s,n) + E\left[e^{-sI(n)} \int_0^{I(n)} (e^{sx} - 1)\, F(dx)\right]. \tag{2.3}$$

Differentiation with respect to s at $s = 0$, or a direct conditional probability argument, now produce equations for $E[I_t]$ and $E[I(n)]$:

$$\frac{d}{dt} E[I_t] = -\lambda E\left[\int_0^{I_t} x\, F(dx)\right] \tag{2.4}$$

and

$$E[I(n+1)] = E[I(n)] - E\left[\int_0^{I(n)} x\, F(dx)\right].$$

In general, no explicit solutions for the expected values are available, but a simple lower bound results from rewriting (2.4) as follows

$$\frac{d}{dt} E[I_t] = -\lambda E\left[I_t \int_0^{I_t} \frac{x}{I_t} F(dx)\right] \tag{2.5}$$

$$\geqslant -\lambda E[I_t F(I_t)]$$

$$\geqslant -\lambda F(I) E[I_t],$$

from which one sees that

$$E[I_t] \geqslant I \exp[-\lambda F(I) t] \tag{2.6}$$

and similarly

$$E[I(n)] \geqslant I[1 - F(I)]^n,$$

so the expected available inventory declines by at most an exponential rate.

2.2. Explicit Solution When the Demand Distribution is Uniform

Although Equation (2.2) seems to be quite intractable for most demand distributions, it can be solved completely when F is uniform:

$$F(x) = \begin{cases} \dfrac{x}{c} & 0 \leqslant x \leqslant c, \\ 1 & c \geqslant x \end{cases}$$

and $c \geqslant I$. In this case (2.2) can be expressed as

$$\frac{\partial \phi}{\partial t} = \lambda E\left[e^{-sI_t} \int_0^{I_t} (e^{sx} - 1) \frac{dx}{c}\right] \tag{2.7}$$

$$= \lambda E\left[\frac{1 - e^{-sI_t}}{sc} - \frac{e^{-sI_t} I_t}{c}\right]$$

$$= \frac{\lambda}{c}\left[\frac{1 - \phi}{s}\right] + \frac{\lambda}{c} \frac{\partial \phi}{\partial s}.$$

In other words, ϕ satisfies a first-order (quasi) linear partial differential equation with initial condition $\phi(s, 0) = e^{-sI}$. Standard procedures (Sneddon [8]) easily yield the solution

$$\frac{1 - \phi(s,t)}{s} = \frac{1 - \exp[-(s + (\lambda/c)t)I]}{s + (\lambda/c)t} \tag{2.8}$$

which gives the desired transform. Passage to the limit as $s \to 0$ in (2.8) shows that

$$E[I_t] = \frac{1 - \exp[-(\lambda t/c)I]}{(\lambda/c)t}. \tag{2.9}$$

This formula can also be derived by first finding an expression for the kth moment of I_t, and then employing a Taylor series argument.

In order to invert the transform in (2.8) note that

$$\int_0^I e^{-sx} P\{I > x\}\, dx = \frac{1 - \phi(s,t)}{s} = \frac{1 - \exp[-(s + (\lambda t/c))I]}{s + (\lambda/c)t} \tag{2.10}$$

which is the transform of a truncated exponential distribution. Thus, by the unicity theorem for Laplace transforms,

$$P\{I > x\} = \begin{cases} \exp[-(\lambda t/c)x] & 0 \leqslant x < I, \\ 0 & I \leqslant x. \end{cases} \tag{2.11}$$

Note that the distribution of I_t is absolutely continuous in the interval $(0, I)$ but that there is a jump at I corresponding to the occurrence of no demand less than, or equal to, I in $(0, t]$:

$$P\{I_t = I\} = \exp[-\lambda t(I/c)]. \tag{2.12}$$

2.3. The Expected Number of Satisfied Demands

Supposing that an initial inventory, or storage capacity, I prevails, it is of interest to compute the probability that a demand is satisfied, and the expected number of demands satisfied in an interval of length t. First notice that if a demand of size $D(t)$ appears at time t, at which moment I_t is available, then

$$P\{D(t) < I_t | I_t\} = F(I_t)$$

is the conditional probability that the demand is satisfied. When F is uniform, as is presently true, we may remove the condition to find that

$$P\{D(t) \leqslant I_t\} = E[F(I_t)] = E\left[\frac{I_t}{c}\right] = \frac{1 - \exp[-(\lambda t/c)I]}{\lambda t}.$$

If $S(t)$ is the number of demands satisfied during the time interval $(0, t]$, then since demands arrive according to a Poisson process with rate λ,

$$E[S(t)] = \lambda \int_0^t E[F(I_u)]\,du = \lambda \int_0^t \frac{1 - \exp[-(\lambda u/c)I]}{\lambda u}\,du \tag{2.13}$$

$$= \gamma + \ln\left[\frac{\lambda tI}{c}\right] + E_1\left[\frac{\lambda tI}{c}\right]$$

where $E_1(\cdot)$ is an exponential integral; Abramowitz and Stegun [1], and $\gamma = 0.5112\ldots$ is Euler's constant.

2.4 The Time of the First Unsatisfied Demand and the Amount of Unused Inventory at that Time

As before F is the common distribution function of the successive demands. Now let τ be the time of the first unsatisfied demand. Then

$$P\{\tau > t | N_t = n\} = P\{D_1 \leqslant I,\ D_2 \leqslant I - D_1,\ \ldots,\ D_n \leqslant I - D_1 - \cdots - D_{n-1}\}$$

$$= F^{(n)}(I)$$

where $F^{(n)}$ denotes the nth convolution of F with itself. Hence,

$$P\{\tau > t\} = \sum_{n=0}^{\infty} e^{-\lambda t} \frac{(\lambda t)^n}{n!} F^{(n)}(I). \tag{2.14}$$

Explicit expressions for the distribution of τ can be obtained in some cases. If F is *uniform* on $[0, c]$ with $c \geqslant I$, then

$$P\{\tau > t\} = e^{-\lambda t} I_0\left[2\left[\frac{\lambda tI}{c}\right]^{1/2}\right] \tag{2.15}$$

where $I_0(z)$ is a modified Bessel function of the first kind of the zeroth order. In this case,

$$E[\tau] = \frac{1}{\lambda} \exp\{I/c\} = \frac{1}{\lambda} \exp\{I/2E[D]\}. \tag{2.16}$$

If F is *exponential* with mean $1/\mu$, then

$$P\{\tau > t\} = \sum_{n=0}^{\infty} e^{-\lambda t} \frac{(\lambda t)^n}{n!} \sum_{k=n}^{\infty} e^{-\mu I} \frac{(\mu I)^k}{k!} \tag{2.17}$$

and

$$E[\tau] = \frac{1}{\lambda} [1 + \mu I] = \frac{1}{\lambda} \left[1 + \frac{I}{E[D]}\right]. \tag{2.18}$$

Note that if I is small relative to $E[D]$, then the expected time to first unsatisfied demand when F is exponential will be greater than the expected time when F is uniform. However, for I large relative to $E[D]$ the expected time for F exponential will be less than the expected time when F is uniform.

Let Y_n be the amount of inventory present at the time of the nth unsatisfied demand. Then for $0 \leqslant a \leqslant I$

$$P\{Y_1 \geqslant I - a\} = \int_0^a R(dy)\, \bar{F}(I - y) \tag{2.19}$$

where

$$R(y) = \sum_{n=0}^{\infty} F^{(n)}(y) \tag{2.20}$$

and

$$\bar{F}(I - y) = 1 - F(I - y). \tag{2.21}$$

Again, explicit expressions for the distribution of Y_1 can be obtained for some distributions F. If F is uniform on $[0, c]$ for $c \geqslant I$, then

$$P\{Y_1 \geqslant I - a\} = 1 - \left(\frac{I - a}{c}\right) \exp\left\{\frac{1}{c} a\right\}. \tag{2.22}$$

If F is a truncated exponential

$$F(x) = \begin{cases} \dfrac{1 - e^{-\mu x}}{1 - e^{-\mu I}} & x \leqslant I, \\ 1 & x \geqslant I, \end{cases} \tag{2.23}$$

then

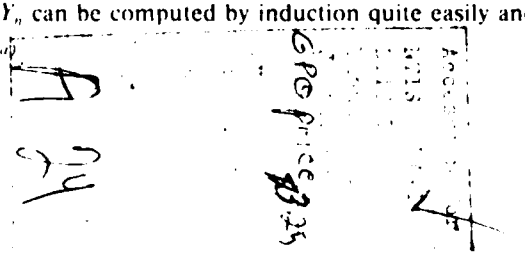
$$P\{Y_1 \geqslant I - a\} = 1 - [e^{\mu a} - e^{\mu I}][1 - e^{\mu I}]^{-1} \exp\{\mu a[1 - e^{\mu I}]^{-1}\}. \tag{2.24}$$

If F has an exponential distribution with mean $1/\mu$, then

$$P\{Y_1 \geqslant I - a\} = e^{\mu(I - a)}. \tag{2.25}$$

In this last case, the distribution function of Y_n can be computed by induction quite easily and

$$P\{Y_n \geqslant I - a\} = e^{n\mu(I - a)} \sum_{k=0}^{n-1} \frac{[\mu(I - a)]^k}{k!}. \tag{2.26}$$



Hence, when F is exponential

$$E[Y_n] = \frac{1}{n\mu}\,[1 - e^{-n\mu t}]. \tag{2.27}$$

In principle, similar results can be obtained for other distributions, but we have found no simple expressions.

2.5. Inventory Costs and Policies

There are at least three monetary quantities which affect the profitability of an inventory policy over a fixed interval of time $(0, t]$: the selling price, p, the storage cost, a, and the cost of lost demands, b. If the storage cost a is charged just on the basis of I (something like warehouse size) then the total expected profit in $(0,t]$ is

$$\begin{aligned} Z(I) &= p(I - E[I_t]) - aI - bS(t) \\ &= (p - a)I - p\left(\frac{\lambda}{c}t\right)^{-1}[1 - \exp[-(\lambda t/c)I]] \\ &\quad - b\left\{\gamma + ln\left[\frac{\lambda t}{c}I\right] + E_1\left(\frac{\lambda t}{c}I\right)\right\} \end{aligned}$$

for the case of uniformly distributed demands; see ((2.9) and (2.13)). One can numerically find the maximum expected profit for this case; nothing explicit seems to be available.

3. THE MANY-BUFFER STORAGE PROBLEM

In this section we will study a model for the situation of Example (c) in Section 1. Messages are successively admitted to the nth buffer until there is a message length that exceeds the remaining capacity of the buffer. The total amount of this message is put in the $(n + 1)st$ buffer and the nth buffer is left forever. Successive messages are then put in the $(n + 1)\,st$ buffer until there is a message whose length exceeds the remaining capacity of the $(n + 1)\,st$ buffer; this message is put in the $(n + 2)nd$ buffer and so on.

Let I denote the common capacity of the buffers and D_i denote the length of message i. Assume $\{D_i\}$ is a sequence of independent identically distributed random variables with distribution F having a density function f such that $f(x) \geq d > 0$ for $x \in [0, I]$. Let $R(x) = \sum_{n=0}^{\infty} F^{(n)}(x)$ be the renewal function associated with F. If $F(I) < 1$, then we will assume that an incoming message to the currently used nth buffer of length greater than I is sent to the $(n + 1)st$ buffer; when it cannot fit into the $(n + 1)st$ buffer, then it is "banished," i.e., sent to some other set of buffers. The next message, however, will try to enter the $(n + 1)st$ buffer. If this message has length greater than I it is banished and the following message will try to enter the $(n + 1)st$ buffer; all messages of length exceeding I will be banished until one appears that is smaller than I and it will be the first entry in buffer $(n + 1)$.

This model has been studied for demand distributions F with $F(I) = 1$ by Coffman et al. [2]. Their approach was to study the Markov process describing the total amount of inventory or space consumed in successive buffers or bins. Here we study the process $\{L_n\}$, where L_n is the size of the demand that first exceeds the remaining capacity of the nth buffer; $\{L_n;\ n = 1, 2, \ldots\}$ is a Markov process. Let

$$K(x, [0,y]) = P\{L_{n+1} \leq y \mid L_n = x\}.$$

Note that

$$P\{L_1 \leqslant y\} = K(0, [0, y])$$

is the same as the sum of the forward and backward recurrence times at time I for a temporal renewal process with interrenewal distribution F; see Feller [5]. Thus for $y \leqslant I$

$$H_1(y) \equiv P\{L_1 \leqslant y\} = \int_{I-y}^{I} R(dz)\,[F(y) - F(I - z)]. \tag{3.1}$$

Note that for $y < I$

$$K(x, [0, y]) = \begin{cases} \int_{I-x-y}^{I-x} R(dz)\,[F(y) - F(I - x - z)] \\ \qquad \text{if } x < I - y; \\ \int_{0}^{I-x} R(dz)\,[F(y) - F(I - x - z)] \\ \qquad \text{if } I - y \leqslant x < I; \\ \int_{I-y}^{I} R(dz)\,[F(y) - F(I - z)] \\ \qquad \text{if } x > I. \end{cases} \tag{3.2}$$

Hence,

$$K(x, dy) = \begin{cases} [R(I - x) - R(I - x - y)]\,F(dy) & \text{if } x < I - y, \\ R(I - x)\,F(dy) & \text{if } I > x > I - y, \\ R(y)\,F(dy) - \int_0^y R(dz)\,f(y - z) + R(dy)F(y) & \\ \qquad \text{if } x = I - y, \\ [R(I) - R(I - y)]\,F(dy) & \text{if } x > I. \end{cases} \tag{3.3}$$

Note that for some $0 < a < b < I$, there exists a $\delta > 0$ such that for all x

$$K^2(x, dy) \geqslant \delta \text{ for } y \in [a, b]$$

where $K^2(x, dy) = \int_0^\infty K(x, dz)\,K(z, dy)$. Hence, hypothesis D' on page 197 of Doob [4] is satisfied. Thus, if

$$K^n(x, A) = P\{L_{1+n} \in A \mid L_1 = x\}$$

for all Borel subsets A, then

$$\lim_{n \to \infty} K^n(x, A) = H(A) \tag{3.4}$$

exists and further the convergence is geometric

$$|K^n(x, A) - H(A)| \leqslant \alpha\gamma^n$$

for some positive constants α and γ, $\gamma < 1$ for all A.

Now let

$$H_n(x) = P\{L_n \in [0, x] \mid L_0 = 0\}.$$

Then a renewal argument can be used to show that for $x \leqslant I$

$$H_{n+1}(x) = \int_{I-x}^{I} H_n * R(dy)\,[F(x) - F(I - y)] + [1 - H_n(I)] \int_{I-x}^{I} R(dy)\,[F(x) - F(I - y)]. \tag{3.5}$$

Taking limits as $n \to \infty$ it is seen that the distribution $H(x)$ satisfies the following equation for

$$H(x) = \int_{I-x}^{I} H * R(dy)\,[F(x) - F(I-y)] + [1 - H(I)] \int_{I-x}^{I} R(dy)\,[F(x) - F(I-y)]. \tag{3.6}$$

Equations (3.1) and (3.6) can be simplified for certain specific distributions F.

3.1 Exponential Demands

For the *exponential* distribution with mean 1 and $x \leqslant I$ the equations are

$$H_1(x) = 1 - e^{-x} - xe^{-x} \tag{3.7}$$

and

$$H(x) = xe^{-x}H(I) + H_1(x) - e^{-x}\int_0^x H(I-x+u)\,du. \tag{3.8}$$

3.2 Uniform Demands

For the uniform distribution on $[0, c]$ with $c \geqslant I$ they simplify to

$$H_1(x) = \exp\left[\frac{1}{c}(I-x)\right] - \left[1 - \frac{x}{c}\right]\exp\left[\frac{1}{c}I\right] \tag{3.9}$$

and

$$\begin{aligned} H(x) = {} & \frac{1}{c}\exp\left[\frac{1}{c}(I-x)\right]\int_0^{I-x}\exp\left[-\frac{1}{c}u\right]H(u)\,du \\ & + \frac{x}{c}H(I) - \frac{1}{c}\exp\left[\frac{1}{c}I\right]\left[1-\frac{x}{c}\right]\int_0^{I}\exp\left[-\frac{1}{c}u\right]H(u)\,du \\ & + [1 - H(I)]\,H_1(x), \end{aligned} \tag{3.10}$$

for $x \leqslant I$. Similar expressions hold for $x > I$, but they are unimportant in the present context.

Equations (3.6), (3.8) and (3.10) do not seem to yield explicit answers. As a result, we have solved (3.8) and (3.10) numerically by iteration using the system of equations

$$H_{n+1}(x) = xe^{-x}H_n(I) + H_1(x) - e^{-x}\int_0^x H_n(I-x+u)\,du \tag{3.11}$$

with H_1 as in (3.7) and

$$\begin{aligned} H_{n+1}(x) = {} & \frac{1}{c}\exp\left[\frac{1}{c}(I-x)\right]\int_0^{I-x}\exp\left[-\frac{1}{c}u\right]H_n(u)\,du \\ & + \frac{x}{c}H_n(I) - \frac{1}{c}\exp\left[\frac{1}{c}I\right]\left[1-\frac{x}{c}\right]\int_0^{I}\exp\left[-\frac{1}{c}u\right]H_n(u)\,du \\ & + [1 - H_n(I)]\,H_1(x) \end{aligned} \tag{3.12}$$

with H_1 as in (3.9). For the cases carried out the convergence is rapid; after $n = 5$ iterations, very little change is noted and convergence has occurred for most practical purposes.

Next let Y_n be the amount of storage space used in the nth bin; the distribution of Y_n is denoted by $G_n(x)$, and

$$G(x) = \lim_{n \to \infty} P\{Y_n \leq x\} = \lim_{n \to \infty} G_n(x)$$

is the long-run distribution. By probabilistic arguments and (3.4)

$$G(x) = \int_0^x H * R(dy)\, \bar{F}(I - y) + [1 - H(I)] \int_0^x R(dy)\, \bar{F}(I - y) \tag{3.13}$$

where $\bar{F}(I - y) = 1 - F(I - y)$ and the long-run average expected capacity of a bin that is actually used is

$$A = \int_0^I x\, G(dx).$$

For the case in which F is exponential with unit mean

$$A = I - [1 - H(I)][1 - e^{-I}] - e^{-I} \int_0^I e^x H(x)\, dx. \tag{3.14}$$

For the case in which F is uniform on $[0, c]$ with $c > I$

$$A = -2 \int_0^I H(u)\, du + \exp\left[\frac{1}{c} I\right] \int_0^I \exp\left[-\frac{1}{c} u\right] H(u)\, du \tag{3.15}$$

$$+ H(I) \left[2I - c \exp\left[\frac{1}{c} I\right] + c\right] + \left[-I + c \exp\left[\frac{1}{c} I\right] - c\right].$$

Numerical solutions were obtained for Equations (3.14) and (3.15) by first computing the probabilities $H_n(x)$, $n = 1, 2, \ldots, 10$ iteratively from (3.7) and (3.11) for the exponential demand case, and from (3.9) and (3.12) for the case of uniform demands. Our technique was simply to discretize x: $x = jh$, $h = I/N$, N being the number of x-values at which $H_n(x)$ is evaluated (values of N from 200-1200 were utilized in order to obtain two-significant digit accuracy). The integrals were then approximated by a summation, i.e. Simpson's rule. Having the values of $H_n(x)$ it is possible to calculate those of $H_{n+1}(x)$, and from these the values of $G_n(x)$ and the mean usage, $E[Y_n]$, may be calculated by numerical integration. In the case of exponential demand very simple upper and lower bounds were obtainable; such bounds were not tight enough to be useful for the uniform case.

The following table summarizes the numerical results. We have compared demand distributions that result, as nearly as possible, in the same probability that an initial demand on an empty bin will be rejected. We have tabulated the expected level to which the bin is filled. It is interesting that the limiting bin occupancy is 0.75 when a uniform demand over the range of the bin size is experienced. This result has been obtained analytically by Coffman et al. [2]; in that paper simple and elegant analytical expressions for G and H also appear for this case. The considerable similarity of the numbers in the rows of the table is notable; apparently the long-run bin occupancy is only slightly larger than is that of the first bin, and the occupancy experienced for uniform demand is only slightly larger than for exponential. Further investigations to examine the reasons for this insensitivity would seem to be of interest.

ACKNOWLEDGMENTS

D. P. Gaver wishes to acknowledge the hospitality of the Statistics Department, University of Dortmund, West Germany, where he was a guest professor during the summer of 1977, and where part of this work was carried out.

Expected Fraction of Bin Filled ($f_n = E[Y_n] \div I$)				
Rejection Probability $\bar{F}(I)$	Exponential Demand f_1	Exponential Demand f_∞	Uniform Demand f_1	Uniform Demand f_∞
0.00	—	—	0.76	0.75
0.05	0.74	0.75	0.74	0.74
0.10	0.69	0.70	0.72	0.72
0.15	0.65	0.66	0.68	0.69
0.20	0.60	0.62	0.64	0.66
0.25	0.56	0.58	0.60	0.62

REFERENCES

[1] Abramowitz, M. and I.A. Stegun, *Handbook of Mathematical Functions.* National Bureau of Standards, AMS 55, Washington, D.C. (1965).

[2] Coffman, E.G., Jr., M. Hofri and K. So, "A Stochastic Model of Bin-Packing," Technical Report, TR-CSL-7811, Computer Systems Laboratory, University of California, Santa Barbara, California (1978), (submitted for publication to a technical journal).

[3] Cohen, J.W., "Single-Server Queues with Restricted Accessibility." Journal of Engineering Mathematics, *3*, 253-284 (1969).

[4] Doob, J.L., *Stochastic Processes*, (John Wiley and Sons, New York, N. Y., 1952).

[5] Feller, W., *An Introduction to Probability Theory and Its Applications, II*, (John Wiley and Sons, New York, N. Y., 1966).

[6] Gavish, B., and P. Schweitzer, "The Markovian Queue with Bounded Waiting Time." Management Science, *23*, 1349-1357 (1977).

[7] Hokstad, P., "A Single-server Queue with Constant Service Time and Restricted Accessibility." Management Science, *25*, 205-208 (1979).

[8] Sneddon, I., *Elements of Partial Differential Equations*, (McGraw-Hill, New York, N. Y. 1957).

RELIABILITY GROWTH OF REPAIRABLE SYSTEMS

Stephen A. Smith and Shmuel S. Oren*

Analysis Research Group
Xerox Palo Alto Research Center
Palo Alto, California

ABSTRACT

This paper considers the problem of modeling the reliability of a repairable system or device that is experiencing reliability improvement. Such a situation arises when system failure modes are gradually being corrected by a test-fix-test-fix procedure, which may include design changes. A dynamic reliability model for this process is discussed and statistical techniques are derived for estimating the model parameters and for testing the goodness-of-fit to observed data. The reliability model analyzed was first proposed as a graphical technique known as Duane plots, but can also be viewed as a nonhomogeneous Poisson process with a particular mean value function.

1. INTRODUCTION

Predicting the reliability of a system or piece of equipment during its development process is an important practical problem. Reliability standards are often a major issue in the development of transportation facilities, military systems, and communication networks. For commercial products that are to be leased and maintained in a competitive marketplace, system reliability estimates strongly influence predicted profitability and customer acceptance. When considering a system that is modified in response to observed failures, most classical statistical estimation techniques are not applicable. This is because the system reliability is improving with time, while most statistical techniques require repeated samples under identical conditions.

A frequently used graphical model of reliability growth of repairable systems is known as "Duane Plots," proposed by J. T. Duane [9]. This model is based on the empirical observation that, for many large systems undergoing a reliability improvement program, a plot of cumulative failure rate versus cumulative test time closely follows a straight line on log-log paper. Several recent papers present applications of Duane plots, e.g., [4], [9] and [10]. Estimating the parameters of the Duane model, i.e., the slope and intercept of the straight line fit, is somewhat difficult to do directly on the graph [5]. Weighted least squares and regression techniques are sometimes used ([9], [10]) to obtain parameter values.

An underlying probabilistic failure model that is consistent with the Duane reliability model is the nonhomogeneous Poisson process (NHPP) whose intensity is total test time raised to some power. (See [7] and [8]). Assuming the sample data consists of all the individual failure times, Crow [7] derived maximum likelihood estimates for the Duane model parameters and a goodness-of-fit test based on the Cramer-von Mises statistic (Parzen [12, p. 143]). A more general NHPP model was proposed by Ascher and Feingold [1], which also used

*Now with Dept. of Engineering Economic Systems, Stanford University, Stanford, CA

the Cramer-von Mises statistic for goodness-of-fit testing. Critical values of this statistic, however, must be obtained by Monte Carlo simulation for each sample size. Crow [7, p. 403] calculated and tabulated values for sample sizes up to sixty. These parameter estimates and goodness-of-fit test deal effectively with Duane model applications having small sample sizes. The facts that all failure times must be stored and the goodness-of-fit measure must be evaluated by simulation make this approach difficult for larger sample sizes. A recent paper by Singpurwalla [13] proposes a time series model for reliability dynamics. This model can, of course, be applied to any type of reliability trend data, but requires data tabulation at a larger number of time stages and does not have the intuitive appeal of the Poisson process for modeling failure occurrences in certain systems.

Our paper develops statistical estimators for the Duane model parameters based on tabulating the number of failures between fixed points in time. This approach has the advantage of using "sufficient statistics" for the data collection, i.e., the dimension of the data does not increase with sample size. Parameter estimates are obtained by maximum likelihood and a goodness-of-fit test based on the Fisher chi-square statistic is derived. This test has the advantage that chi-square tables are readily available for all sample sizes and significance levels. The accuracy of the chi-square test decreases, however, as the sample size gets small. Sample sizes for which the techniques of this paper apply are found in developmental systems that experience frequent, minor failures such as paper jams in photo copy machines, voltage fluctuations in power supply systems, faults in semiconductor manufacturing processes, etc. The last section of this paper illustrates the application of the estimation and goodness-of-fit techniques to a representative set of simulated failure data.

Regardless of how the parameters of the Duane model are obtained, considerable caution is required when extrapolating reliability trends beyond the observed data to future time points. Major breakthroughs or setbacks in the reliability improvement program may cause significant deviations from the straight line projections. Some users recommend reinitializing the model and shifting to a new straight line fit when major changes in the program occur. Even if one is uneasy about extrapolating the reliability growth model to estimate future reliability, it remains a valuable tool for obtaining a "smoothed" estimate of current system reliability. While reliability is changing, sample sizes at any point in time are not sufficient for conventional statistical estimation techniques. With a dynamic reliability model, past and current failure data can be combined to obtain estimates of current reliability based on fitting all observed data.

2. THE DUANE MODEL

The Duane model states that cumulative failure rate versus cumulative test time, when plotted on log-log paper, follows approximately a straight line. More precisely, if we let $N(0,t)$ represent the total number of failures observed up to time t, we have that

$$\log[N(0,t)/t] \approx -b \log t + a, \tag{2.1}$$

where the fitted parameters are $a, b > 0$. The relationship is meaningless at $t = 0$ but, as most users point out ([5],[9]), a certain amount of early data is generally excluded from the fit because it is influenced by factors such as training of personnel, changes in test procedures, etc. Equation (2.1) therefore implies that

$$N(0,t)/t \approx \alpha t^{-b}, \text{ where } \alpha = \log a,$$

for t beyond a certain point. It should be emphasized that, in all applications, time t corresponds to cumulative operating time or test time. For the results of this paper it is most convenient to write the Duane model as:

$$N(0,t) \approx \alpha t^{\beta}, \text{ where } \beta = 1 - b. \tag{2.2}$$

For a fairly diverse set of observed systems, Codier [5, p. 460] has found b to be generally between 0.3 and 0.5, corresponding to β between 0.5 and 0.7.

3. AN UNDERLYING STATISTICAL MODEL

In this section we describe a statistical model for the failure process that is consistent with assuming that the observed failure data fits the Duane model. Suppose the probability that the system fails at time t (strictly speaking in a small interval $[t, t + dt)$), regardless of the past, is determined by a hazard function $h(t)$. That is,

$$h(t)dt \approx P\{\text{the system fails in the interval}[t, t + dt)\},$$

independent of its previous failure and repair history. The expected number of failures in any time interval $[t_1, t_2)$ of operating time is then given by the mean value function

$$M(t_1, t_2) = \int_{t_1}^{t_2} h(t)\,dt. \tag{3.1}$$

Furthermore, it can be shown (See Parzen [12, Sect. 4.2]) that $N(t_1, t_2)$, the number of failures observed in some future time interval $[t_1, t_2)$, has probability distribution

$$P\{N(t_1, t_2) = k\} = ([M(t_1, t_2)]^k/k!) \exp\{-M(t_1, t_2)\} \quad k = 0, 1, 2, \ldots . \tag{3.2}$$

In addition to its mathematical convenience, this model has considerable intuitive appeal. The simple Poisson process has been used successfully to model the failure occurrences of many devices, or collections of devices operating in series. One may think of a system having a nonhomogeneous Poisson failure process as a large collection of simpler devices in series, with individual device failure modes being gradually removed with time.

The mean value function

$$M(t_1, t_2) = \alpha(t_2^{\beta} - t_1^{\beta}), \quad \text{where } \alpha, \beta > 0; \tag{3.3}$$

corresponding to $h(t) = \alpha\beta t^{\beta - 1}$, is of particular interest. Crow [7, p. 405] pointed out that the number of failures from a process with this mean value function will approximate the Duane Model by observing that

$$\log[M(t)/t] = \log \alpha + (\beta - 1) \log t, \quad \text{where } M(t) = M(0, t).$$

This means that system failure data from a NHPP with mean value function $M(t)$ will approach the Duane model with probability one. Conversely, this process with mean value function $M(t)$ is the only model with independent increments that approximates the Duane model in a probabilistic sense for sufficiently large sample sizes. We will not give a proof of these statements but refer the reader to Parzen [12, ch. 4] or Donelson [8] for a complete discussion.

4. SELECTING A STARTING TIME

The Duane reliability model and the expected number of failures in Equation (3.3) are both nonlinearly dependent on the choice of the time origin. That is, if we begin observing failures at time $t = t_0 > 0$ and ignore the first $N(0, t_0)$ failures and the time interval $[0, t_0)$, we do not obtain the same parameters α and β by fitting the subsequent data. Since the logarithm is a strictly concave function, there is only one choice of t_0 that can give a straight line fit to the data on log-log paper. Specifying the operating time t_0 that is assumed to have elapsed before the beginning of the modeling process is therefore an important step.

Some users of the Duane Model ([5],[10]) suggest reducing the cumulative failures and observation time by removing early data to obtain a straight line fit. This is done graphically by

[illegible] plotting the data until a straight line fit is [illegible] shape of the graph of cumulative failures versus [illegible] downward bending (concave), so it is not hard to [illegible]

[illegible] line trend on log-log paper is observed before the noisy [illegible] is shifted further to the right, the straight line shape will [illegible] most that can be said in this case is that, for t greater than [illegible] the model. The statistical model (3.3) can still be applied, however, [illegible] number of failures $N(t_1,t)$ after the first $N(0,t_1)$ fit the NHPP with [illegible] for [illegible]

5. ESTIMATING THE MODEL PARAMETERS

If the Duane model is applied graphically, the user can attempt to estimate the parameters α and β by drawing the best straight line through the plotted points. This is somewhat tricky because, with cumulative failure data, the later points should be weighted more heavily in determining the fit. This section describes a statistical estimation procedure based on the NHPP model of the failure process. We consider two possibilities for collecting and recording system failure times. The first is to record the occurrence time of each failure, which yields a sequence of observed times $T_1, T_2, \ldots, T_N$. This case has been analyzed by Crow in [7] and the maximum likelihood estimates are given by

$$\alpha^* = N/T_N^{\beta^*} \tag{5.1}$$

and

$$\beta^* = -N/\sum_{i=1}^{N} \log(T_i/T_N). \tag{5.2}$$

A goodness-of-fit test corresponding to these estimators is derived in [7] and critical values of the error statistic are tabulated for sample sizes 2-60.

If large numbers of failures are observed, it is often convenient to record only the aggregate number of failures between each pair in a sequence of fixed time points $t_0, t_1, \ldots, t_n$. In this case the data is in the form $N_1, N_2, \ldots, N_n$, where N_i = number of failures observed in the interval (t_{i-1}, t_i). Maximum likelihood estimates and a goodness-of-fit criterion for observations in this form are developed in the next few paragraphs.

Maximum Likelihood Estimates for the Aggregated Case

We first calculate the likelihood function for the data $N_1, N_2, \ldots, N_n$, given the time points $t_0, t_1, \ldots, t_n$ and the assumed form of the mean value function in Equation (3.3). The probability of N_i system failures in the interval (t_{i-1}, t_i) is obtained from Equation (3.2). Since the underlying model assumes that each of the time segments is independent, the likelihood function can be written as a product of these probabilities,

$$L(\alpha,\beta) = \prod_{i=1}^{n} P\{N(t_{i-1}, t_i) = N_i\} = \exp\{-M(t_0,t_n)\} \prod_{i=1}^{n} ([M(t_{i-1}, t_i)]^{N_i}/N_i!). \tag{5.3}$$

To simplify the calculation of the estimators, we take the log of $L(\alpha,\beta)$, noting that maximizing the log will yield the same maximum likelihood estimates. From (5.3) we have

$$\log L(\alpha,\beta) = -\alpha(t_n^\beta - t_0^\beta) + \sum_{i=1}^{n} N_i[\log\alpha + \log(t_i^\beta - t_{i-1}^\beta)] - \sum_{i=1}^{n} \log N_i!. \tag{5.4}$$

Taking the partial derivatives $(\partial \log L)/\partial\alpha = 0$ and $(\partial \log L)/\partial\beta = 0$, we obtain the equations for the maximum likelihood estimates,

$$\alpha^* = N/(t_n^{\beta^*} - t_0^{\beta^*}), \quad \text{where } N = \sum_{i=1}^{n} N_i \tag{5.5}$$

$$0 = \sum_{i=1}^{n} N_i \left[\frac{\log t_i - \rho_i \log t_{i-1}}{1 - \rho_i} - \frac{\log t_n - \rho_0 \log t_0}{1 - \rho_0} \right], \tag{5.6}$$

where

$$\rho_i = (t_{i-1}/t_i)^{\beta^*}, \ i = 1, 2, \ldots, n, \text{ and } \rho_0 = (t_0/t_n)^{\beta^*}.$$

Equation (5.6) is an implicit function of β^*, but can be solved iteratively by a computer algorithm or programmable calculator, because the right hand side is strictly decreasing in β^*. To verify this fact, consider any two times t, t' and compute the derivative

$$(\partial/\partial\beta)[\log t - T^\beta \log t']/[1 - T^\beta] = -T^\beta (\log T)^2/(1 - T^\beta)^2, \quad \text{where } T = t/t'. \tag{5.7}$$

This derivative is negative and decreasing in T for $0 < T, \beta < 1$. The derivative of the sum in (5.6) is a sum of terms involving the difference of the derivative (5.7) evaluated at $T = t_{i-1}/t$ and t_0/t_n. The fact that (5.6) is decreasing in β^* follows from the fact that (5.7) is decreasing in T, i.e., its largest or least negative value occurs at $T = t_0/t_n$. Therefore, (5.6) has a unique solution.

6. GOODNESS-OF-FIT CRITERION

This section describes a procedure for testing the goodness-of-fit of the observed failure data to the NHPP. We assume that the parameters α^* and β^* are obtained from the maximum likelihood estimates (5.5) and (5.6). From the form of (5.5), it is clear that the estimate α^* is defined in such a way that the total number of observed failures N always equals the expected number of failures for the time period (t_0, t_n). That is, α^* is defined so that

$$N = E\{N | \alpha^*, \beta^*\} = \alpha^*(t_n^{\beta^*} - t_0^{\beta^*}).$$

Therefore, there is no difference between the observed versus predicted *total* number of failures. The goodness-of-fit measure must therefore be based on the differences between the observed incremental failures $N_1, N_2, \ldots, N_n$, and the predicted values

$$E\{N_i | \alpha^*, \beta^*\} = \alpha^*(t_i^{\beta^*} - t_{i-1}^{\beta^*}), \quad i = 1, 2, \ldots, n. \tag{6.1}$$

Assuming the estimate (5.5) is used for α^*, the likelihood function for a goodness-of-fit statistic will be expressed only in terms of β^*. Since the NHPP has independent increments, the probability that a given failure occurs in the interval (t_{i-1}, t_i) is the expected number of failures for that interval, divided by the total number of failures. This is written as

$$p_i = p_i(\beta^*) = [\alpha^*(t_i^{\beta^*} - t_{i-1}^{\beta^*})]/[\alpha^*(t_n^{\beta^*} - t_0^{\beta^*})], \quad i = 1, 2, \ldots, n, \tag{6.2}$$

where the α^* parameter obviously cancels out. The likelihood function for a set of observed failures $N_1, N_2, \ldots, N_n$, given N, is therefore the multinomial

$$\binom{N}{N_1, N_2, \ldots, N_n} p_1^{N_1} p_2^{N_2} \cdots p_n^{N_n}, \quad \text{where } N_1 + N_2 + \cdots + N_n = N,$$

which depends only on β^*. The parameter α^* can be regarded as a scale parameter that guarantees the model will fit the total number observed of failures N.

We now show how the goodness-of-fit of the incremental failure data can be measured by the Fisher chi-square statistic

$$\chi^2 = \sum_{i=1}^{n} (N_i - Np_i)^2/Np_i. \tag{6.4}$$

The use of this statistic as a goodness-of-fit measure is based on the following theorem, which has been restated in the context of this discussion.

THEOREM 6.1: Let the parameters $p_1, p_2, \ldots p_n$, with $\Sigma p_i = 1$, be functions of a parameter β and let a particular value β' be determined from

$$0 = \sum_{i=1}^{n} (N_i/p_i)(\partial p_i/\partial \beta)\Big|_{\beta=\beta'}. \tag{6.5}$$

Then the statistic (6.4) with $p_i = p_i(\beta')$, $i = 1, 2, \ldots, n$, has approximately a chi-square distribution with $n - 1$ degrees of freedom ($\chi^2(n-1)$) for large N. The proof of this result is quite lengthy and can be found in [6, pp. 424-434].

To apply this result to our particular problem, we must show that β' equals the estimator β^* defined by Equation (5.6). Using $p_i(\beta)$ as defined in Equation (6.2), and differentiating with respect to β, one can verify that Equation (6.5) reduces to Equation (5.6). Thus, $\beta' = \beta^*$ and, since (5.6) has only one solution, the value is unique.

The chi-square error statistic (6.4) has an additional intuitive interpretation for this application. Suppose α and β are the "true" parameters of the underlying nohomogeneous Poisson process, i.e., the values to which the estimators α^* and β^* must eventually converge for very large sample sizes. Then the "true" variance of the number of observed failures in $[t_{i-1}, t_i)$, i.e., the limiting value for the sample variance of a large number of observations, is given by

$$\mathrm{Var}\{N_i|\alpha, \beta\} = \alpha(t_i^{\beta} - t_i^{\beta}) \quad i = 1, 2, \ldots, n.$$

Consider $$W(\alpha^*, \beta^*) = \sum_{i=1}^{n} (N_i - E\{N_i|\alpha^*, \beta^*\})^2/\mathrm{Var}\{N_i|\alpha, \beta\},$$

which is the sum of square errors between the observed and estimated failures, weighted by the true variance for each of the time intervals. Suppose we minimize this with respect to α^* and β^* by solving $(\partial W/\partial \alpha^*) = 0$ and $(\partial W/\partial \beta^*) = 0$. If we then substitute our "best estimates", α^* for α and β^* for β, these two equations reduce to the maximum likelihood equations, (5.5) and (5.6), respectively. Birnbaum [2, p. 251-2] also points that if we minimize the chi-square statistic (6.4) with respect to β, the estimate obtained must approach the estimate β' that satisfies (6.5) as the sample size approaches infinity.

This goodness-of-fit criterion measures, in effect, how well the observed data fits a NHPP with mean value function $M(t)$, where β^* is the "best" growth parameter for the observed data. If the $\chi^2(n-1)$ statistic (6.4) exceeds the critical value at a reasonable significance level, such as 0.05 or 0.1, the model should be rejected. Since Theorem 6.1 gives only an asymptotic result, it is important to discuss the sample size requirements for applying it. Given the popularity of this test, there has been considerable experience with various types of data. A common criterion is that N and, in this case the time points $t_0, t_1, \ldots, t_n$, must be such that $Np_i \geq 10$ for all i. (See Birnbaum [2, p. 248]).

7. APPLICATION EXAMPLE

As an illustration, we will determine the estimators α^* and β^* and apply the goodness-of-fit test to the sample data in Table 1. We assume that the failures of the system were only monitored at fixed points of time so that the observed data consists of the first two columns of the table. These data points were generated by computer simulation with failures sampled from a NHPP with mean value function $M(t)$, having parameters $\alpha = 10.0$, $\beta = 0.5$. Failure times $T_1, T_2, \ldots$ from this distribution can be generated sequentially from a set of random samples $U_1, U_2, \ldots$ from the uniform distribution by means of the transformation

$$T_{i+1} = [T_i^\beta - (1/\alpha) \log U_{i+1}]^{1/\beta}, \quad T_0 = 0, \quad i = 0, 1, 2, \ldots \quad . \tag{7.1}$$

TABLE 1

	Time Interval	Observed Failures	Predicted Failures	Standard Deviation	Normalized Error
1	400 - 800	63	78	8.8	2.88
2	800 - 1200	63	61	7.8	0.07
3	1200 - 1600	54	51	7.1	0.18
4	1600 - 2000	51	46	6.8	0.54
5	2000 - 2500	68	51	7.1	5.67
6	2500 - 3000	49	46	6.8	0.20
7	3000 - 3500	34	43	6.6	1.88
8	3500 - 4000	39	40	6.3	0.03
9	4000 - 4500	39	38	6.2	0.02
10	4500 - 5000	43	36	6.0	1.36
11	5000 - 5500	39	34	5.8	0.74
12	5500 - 6000	36	33	5.7	0.27
13	6000 - 6500	28	31	5.6	0.29
14	6500 - 7000	22	30	5.5	2.13
15	7000 - 7500	35	29	5.4	1.24
16	7500 - 8000	32	28	5.3	0.57
17	8000 - 8500	22	27	5.2	0.93
18	8500 - 9000	19	27	5.2	2.37
19	9000 - 9500	19	26	5.1	1.88
					23.25

The data in Table 1 was used to obtain maximum likelihood estimates α^* and β^* from Equations (5.5) and (5.6). This was done by calculating various values of the right hand side of (5.6) as a function of β until the minimizing value β^* was determined to two decimal places. This gave $\beta^* = 0.52$ and $\alpha^* = 7.97$, where α^* was determined from (5.5) with $\beta^* = 0.52$.

The accuracy of β^* is reasonably close to the correct value $\beta = 0.5$, but the estimate of α^* is off by more than 20%. Other calculations with different sets of random numbers produced errors in both directions but generally resulted in an α^* error several times larger than the β^* error, on a percentage basis. This seems to indicate that one is more likely to estimate slopes of the Duane Plot lines accurately than to estimate the intercepts accurately with the maximum likelihood estimates. Naturally, as the number of observation points in Table 1 is increased, the estimates become more accurate. Accuracy was not improved much by increasing the number of time points from 20, as shown in the table, to 100 and the sign of the error for a given example generally did not change as the number of observation points was increased.

while holding the underlying failure points fixed. Bringing the estimate α^* to within 5% of the correct value typically required 300 to 500 observation time points for the computed examples.

To illustrate the use of the goodness-of-fit test we calculate the chi-square statistic (6.4) for this table. The "Predicted Failures" between the various time points are given by

$$Np_i = \alpha^*(t_{i+1}^{\beta^*} - t_i^{\beta^*}), \quad i = 1, 2, \ldots, 19.$$

The normalized error terms as in (6.4) are given by

$$(N - Np_i)^2/(Np_i).$$

The sum of these errors, when compared with a $\chi^2(18)$ error table, is less than the critical values 25.99 and 28.87, associated with significance levels 0.1 and 0.05, respectively.

For many applications of the model it is more important to predict the number of failures that will occur in the next time period than to obtain accurate estimates for α and β. In such cases the estimators obtained from 10-20 time points appear to be sufficiently accurate. This is because there is a range of α, β pairs that provide almost as good a fit to the observed data as the optimal ones and any parameters in this range provide a satisfactory predictive model.

To illustrate the prediction accuracy of the estimates $\beta^* = 0.52$, $\alpha^* = 7.97$ obtained from Table 1, we generated simulated failures out to 40,000 time units. The number of failures predicted by extrapolating with the estimated parameters and with the true parameters are compared in Table 2. The errors in predicting failures caused by inaccuracy in estimating the parameters is much less than the random errors that occur due to stochastic variations of the failure process. This was found to be the case in several similar experiments.

TABLE 2

Time Interval	Simulated Failures	Estimated Extrapolation	True Extrapolation	Standard Deviation
9500 - 10,000	24	25	25	5.0
9500 - 15,000	235	251	250	15.8
9500 - 20,000	412	443	439	21.0
9500 - 30,000	715	766	757	27.5
9500 - 40,000	999	1041	1025	32.0

8. CONCLUSION

Choosing the fixed time points between which to tabulate failures is mainly a question of engineering judgement. The time points might be selected, for example, to correspond to milestones in the reliability development program. The parameter estimates and goodness-of-fit tests obtained in this paper and those obtained by Crow are essentially complementary with respect to various applications of the Duane model. It is not possible to determine the precise sample size at which one approach becomes more advantageous than the other. Based on experience, the chi-square goodness-of-fit test tends to reject most sample data, including data that fits the model, when sample sizes are too small. Therefore, rejection of the model by the chi-square test, based on data with a questionable total number of samples, might be viewed as inconclusive and the more accurate test developed by Crow could then be applied. For large sample sizes that have at least 10 failures between time points, the chi-square test should be accurate and is computationally easier. Data that fails to fit the NHPP model with mean value

function $M(t)$ based on these tests requires a more general approach. A NHPP model with a different intensity such as discussed in [1], or a less constrained model such as [13] might then be tested.

BIBLIOGRAPHY

[1] Ascher, H. and H. Feingold, "'Bad as Old' Analysis of System Failure Data," *Proceedings of the Eighth Reliability and Maintainability Conference* (July 1969).

[2] Birnbaum, Z.W., *Introduction to Probability and Mathematical Statistics* (Harper and Brothers, New York, N. Y., 1962).

[3] Barlow, R.E. and F. Proschan, *Statistical Theory of Reliability and Life Testing* (Holt, Rinehart and Winston, Inc., 1975).

[4] Chapman, W.A. and D.E. Beachler, "Reliability Proving for Commercial Products," *Proceedings 1977 Annual Reliability and Maintainability Symposium.*

[5] Codier, E.O., "Reliability Growth in Real Life," *Proceedings Annual Symposium on Reliability* (1968).

[6] Cramer, H. *Mathematical Methods of Statistics*, (Princeton University Press, Princeton, New Jersey, 1946).

[7] Crow, L.H., "Reliability Analysis for Complex Repairable Systems," *Reliability and Biometry: Statistical Analysis of Lifelength*, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, Pennsylvania (1974).

[8] Donelson, J., "Duane's Reliability Growth Model as a Nonhomogeneous Poisson Process," Institute for Defense Analysis paper P-1162 (April 1975).

[9] Duane, J.T., "Learning Curve Approach to Reliability Monitoring," *IEEE Transactions on Aerospace*, *2* (1964).

[10] Hovis, J.B., "Effectiveness of Reliability System Testing on Quality and Reliability," *Proceedings 1977 Annual Reliability and Maintainability Symposium.*

[11] Mead, P.H., "Duane Growth Model: Estimation of Final MTBF with Confidence Limits Using a Hand Calculator," *Proceedings 1977 Annual Reliability and Maintainability Symposium.*

[12] Parzen, E., *Stochastic Processes* (Holden Day, 1962).

[13] Singpurwalla, N.D., "Estimating Reliability Growth (or Deterioration) Using Time Series Analysis," Naval Research Logistics Quarterly, 25, 1-14 (1978).

ON THE DISTRIBUTION OF THE OPTIMAL VALUE FOR A CLASS OF STOCHASTIC GEOMETRIC PROGRAMS*

Paul M. Ellner† and Robert M. Stark

Department of Mathematical Sciences
University of Delaware
Newark, Delaware

ABSTRACT

An approach is presented for obtaining the moments and distribution of the optimal value for a class of prototype stochastic geometric programs with lognormally distributed cost coefficients. It is assumed for each set of values taken on by the cost coefficients that the resulting deterministic primal program is superconsistent and soluble. It is also required that the corresponding dual program has a unique optimal point with all positive components. It is indicated how one can apply the results obtained under the above assumptions to stochastic programs whose corresponding deterministic dual programs need not satisfy the above-mentioned uniqueness and positivity requirements.

1. INTRODUCTION

This paper is concerned with deriving the distribution and/or moments of the optimal value for a class of stochastic prototype geometric programs in which a subset of the cost coefficients are lognormally distributed. The programs are assumed to be superconsistent and soluble for all positive values that can be taken on by the cost coefficients. It is also required that the dual of a program has a unique optimal point, δ_c, with all positive components, for all possible values that can be taken on by the components of the cost vector c. Such programs include soluble programs with no forced constraints. Also included are superconsistent soluble programs whose forced constraints are nonredundant (and hence active at optimality) and whose forced constraint gradients are linearly independent at optimality, for each positive-valued cost vector c.

The class of problems specified above, while of interest in themselves, can be used to obtain the distribution and/or moments of the optimal value for more general classes of stochastic prototype geometric programs. This will be indicated in Section 6.

The distribution and/or moments of the optimal value of a stochastic program will be expressed in terms of the density function of a vector $\tilde{L} \underline{\underline{\Delta}} (\log K_0, \log K_1, \ldots, \log K_d)'$, where log denotes the natural logarithm function, d is the degree of difficulty of the program, and

*This research was supported in part by the Office of Naval Research Contract N00014-75-C-0254
†Now with U.S. Army Materiel Systems Analysis Activity, Aberdeen Proving Ground, Maryland

$$\log A_j = \sum_{i=1}^{n} b^{(j)} \log c_i \quad \text{for } j \in \{0, 1, \ldots, d\}.$$

In the above $(c_1, \ldots, c_n)'$ is the vector of cost coefficients (where "$'$" denotes transpose) and the $b^{(j)}$ are constants that are independent of the c_i.

One advantage to deriving the distribution and/or moments of the optimal value in terms of the density function of $\tilde{L}$ is that the vector $\tilde{Y}$ is normally distributed when the stochastic c_i are jointly lognormally distributed. Furthermore, under certain conditions, it is reasonable to expect that $\tilde{L}$ behaves approximately as if the vector of stochastic cost coefficients were lognormally distributed even when it is not. More precisely, if the stochastic cost coefficients, $\{c_i | i \in I\}$, are positive-valued and the variates $\{\log c_i | i \in I\}$ are independent with finite means, variances, and third order absolute central moments, then one can apply a central limit theorem for random vectors to the relation $\tilde{L} = \sum_{i=1}^{n} \tilde{Z}^{(i)}$, where $\tilde{Z}^{(i)} \triangleq (b^{(0)i} \log c_i, b^{(1)i} \log c_i, \ldots, b^{(d)i} \log c_i)'$ [11]. Thus, under the above conditions, one might expect that $\tilde{L}$ tends to be normally distributed *provided the stochastic c_i are positive-valued*, strictly unimodal, continuous variates; the number of indices in I is "large" in comparison to $d + 1$, and no partial sum of $d + 1$ of the $\tilde{Z}^{(i)}$ is "excessively" dominant in the sum for $\tilde{L}$.

The results of this paper should be of interest in instances where the operating or construction costs associated with a contemplated project or engineering system can be adequately approximated as the optimal value of a stochastic prototype geometric program with lognormally distributed cost coefficients. In such cases a knowledge of the distribution function and/or moments would be useful as a predictive tool in financial planning. For instance, if the distribution function of the optimal value were known one would be able to predict with a given probability that a proposed system's operating or construction costs incurred over a given period would lie within a specified set of limits.

To reflect the uncertainty as to the future costs, c_i, that will be encountered during the construction or operating period of interest a cost analyst often subjectively chooses a distribution for each cost c_i. Cost analysts have frequently found families of positive-valued random variables that are continuous and strictly unimodal useful for this purpose [9]. The lognormal random variables form a two parameter family that meets these specifications. Recall a random variable X is said to be lognormal iff $\log X$ is normally distributed. Properties of lognormal random variables can be found in [2].

Cost analysts are most often concerned with the distribution of values of c_i about a central value and not with tail values. Thus, an analyst who wishes to utilize the present results might proceed to express his uncertainty about the future value of cost coefficient c_i as follows:

1. Assume c_i is lognormally distributed and subjectively choose the median value of c_i, denoted by ξ_i.

2. Specify an interval of interest about ξ_i of the form $(\theta_i^{-1}\xi_i, \theta_i\xi_i)$ where $\theta_i \in (1, \infty)$.

3. Subjectively choose $\delta_i \in (0, 1)$ such that $1 - \delta_i$ reflects one's belief that $c_i \in (\theta_i^{-1}\xi_i, \theta_i\xi_i)$, i.e., the more confident one is that $c_i \in (\theta_i^{-1}\xi_i, \theta_i\xi_i)$ the closer $1 - \delta_i$ should be chosen to 1.

4. Calculate the unique value of the standard deviation of c_i that is consistent with (1) and the equation $Pr(\theta_i^{-1}\xi_i < c_i < \vartheta_i\xi_i) = 1 - \delta_i$ where Pr denotes the probability function associated with c_i.

Results of the paper do not require that the stochastic c_i be independently distributed. Thus, for every pair of stochastic cost coefficients c_i, c_j $(i \neq j)$ the analyst may subjectively choose a number between -1 and 1, the correlation coefficient ρ_{ij} of log c_i and log c_j, to reflect his opinion as to the interdependency of c_i and c_j. The theory allows for the possibility that $\rho_{ij} = \pm 1$ (i.e., with probability 1 $c_j = \alpha c_i^{\beta}$ for some constants $\alpha \in (0, \infty)$ and $\beta \in (-\infty, \infty)$).

In Section 2 the notation used in connection with the deterministic and stochastic geometric programming problem and its dual and transformed dual is presented. Also the special role of the transformed dual program in obtaining the distribution and/or moments of the optimal value of the primal program is indicated.

Section 3 presents and discusses the assumptions placed upon the primal program throughout Sections 3 through 5 and the appendices. Additionally, useful properties of the density functions of $\tilde{L}$ and $L \underline{\Delta} (\log K_1, \ldots, \log K_d)'$ are stated.

In Section 4 we use the density functions of L and $\tilde{L}$, together with the maximizing equations for an unconstrained transformed dual program, to obtain the density functions of r and $(r, v(P_c))$. Here r denotes the random vector of the optimal point of the unconstrained stochastic transformed dual program and $v(P_c)$ denotes the optimal value of the stochastic primal program. We then obtain the density function of $v(P_c)$ as a marginal density of $(r, v(P_c))$.

In Section 5 we use the density function of r to derive a formula that expresses each moment of $v(P_c)$ as the integral of an explicitly given integrand over an explicitly specified convex polyhedral subset of R^d, where d is the degree of difficulty of the stochastic primal program.

Section 6 briefly indicates how the preceding results can be used to calculate the distribution and/or moments of $v(P_c)$ when P_c need not satisfy all the assumptions of Section 3.

Appendix A contains the statement and proof of a lemma from which important properties of $\tilde{L}$ and L immediately follow. These properties are stated in Theorem 1 of Section 3.

Finally, in Appendix B we establish that boundedness of the dual feasible set is a sufficient condition for the existence of all the moments of $v(P_c)$, under the assumptions of Section 3.

2. NOTATION AND PRELIMINARIES

We shall now review the definitions and notation used in connection with prototype geometric programming that will be utilized in the paper. In the following, for every positive integer ν, $<\nu> \underline{\Delta} \{1, \ldots, \nu\}$ and $<\bar{\nu}> \underline{\Delta} \{0, 1, \ldots, \nu\}$. Also, for every matrix P, P' denotes the transpose of P. All elements of Euclidean n-space, R^n, will be viewed as column vectors of n real numbers and the zero vector will be denoted by $\underline{0}$.

Recall a prototype primal geometric program has the following form [4]: inf $g_0(t)$ subject to $g_\kappa(t) \leqslant 1$ $\forall$ $\kappa \in <p>$ and $t_i > 0$ $\forall$ $i \in <m>$ where $t = (t_1, \ldots, t_m)'$ and $g_\kappa(t) \underline{\Delta} \sum_{i \in J_\kappa} c_i$

$\prod_{i=1}^{m} t_i^{a_{ij}}$ for $\kappa \in <p>$. In the above $A = (a_{ij})$ is an n by m matrix with real entries called the exponent matrix and $c = (c_1, \ldots, c_n)'$ is a vector of positive numbers called the vector of cost coefficients. Also, $J_\kappa \triangleq \{m_\kappa, m_{\kappa+1}, \ldots, n_\kappa\}$ where $m_0 = 1$, $m_\kappa = n_{\kappa-1} + 1$ for $\kappa \in <p>$, and $n_p = n$. The constraints $g_\kappa(t) \leqslant 1$ are called forced constraints and we allow the possibility that a primal program has no forced constraints.

In this paper we shall be concerned with problems of the above form when some or all of the cost coefficients are stochastic variables that are lognormally distributed. Thus, we shall assume there exists $I \subset <n>$ such that $I \neq \phi$ and $i \in I$ iff c_i is stochastic. Let $c_I \triangleq (c_{i_1}, \ldots, c_{i_\omega})'$ where $i_1 < \ldots < i_\omega$ and $I = \{i_\nu \mid \nu \in <\omega>\}$. Thus, c_I is a random vector formed from the stochastic cost coefficients. Values taken on by c_I will be denoted by $\bar{c}_I$. Also $\bar{c}$ will denote the value taken on by cost coefficient vector c when c_I takes on the value $\bar{c}_I$. We shall let P_c and $P_{\bar{c}}$ denote the corresponding stochastic and deterministic prototype primal geometric programs. Furthermore, $v(P_{\bar{c}})$ will denote the optimal value of $P_{\bar{c}}$ and $v(P_c)$ will denote the stochastic variable that takes on the value $v(P_{\bar{c}})$ when c_I takes on the value $\bar{c}_I$.

The stochastic program P_c is not convenient to work with due to possible randomness in coefficients of the forced constraints. To find computationally tractable bounds on the solution of a two stage geometric program with stochastic cost coefficients, Avriel and Wilde [3] considered the stochastic problem D_c in place of P_c where, for every $\bar{c} > \underline{0}$, $D_{\bar{c}}$ is the dual of $P_{\bar{c}}$ as given in [4]. The stochastic program D_c has the attractive feature that all its randomness is confined to the objective function. To see this recall D_c is the following program: sup $\prod_{i=1}^{n} (\bar{c}_i/\delta_i)^{\delta_i} \prod_{\kappa=1}^{p} \lambda_\kappa(\delta)^{\lambda_\kappa(\delta)}$ subject to the normality condition $\sum_{i=1}^{n_0} \delta_i = 1$, *the orthogonality conditions* $\sum_{i=1}^{n} a_{ij}\delta_i = 0$ for $j \in <m>$, and the positivity conditions $\delta_i \geqslant 0$ for $i \in <n>$. In the above, for every $\kappa \in <p>$, $\lambda_\kappa(\delta) \triangleq \sum_{i \in J_\kappa} \delta_i$ for $\delta \in R^n$. Also, in evaluating the dual objective function one uses the convention that $x^y = x^{-y} = 1$ for $x = 0$. When P_c has no forced constraints we set $p = 0$ and define the expression $\prod_{\kappa=1}^{p} \lambda_\kappa(\delta)^{\lambda_\kappa(\delta)}$ to be 1.

Under rather general conditions one has $v(D_c) = v(P_{\bar{c}})$ for $\bar{c} \in R^n_+$ [4, Ch. 6] (where R^n_+ denotes the positive orthant of R^n and $v(D_{\bar{c}})$ denotes the optimal value of $D_{\bar{c}}$). This is true, e.g., if $P_{\bar{c}}$ is superconsistent and soluble [4, Ch. 4]. Thus, frequently the distribution function of $v(D_c)$ will be the same as that for $v(P_c)$, where $v(D_c)$ denotes the stochastic variable that takes on the value $v(D_{\bar{c}})$ when c_I takes on the value $\bar{c}_I$. Obtaining the distribution function and/or moments of $v(D_c)$ is facilitated by the fact that the constraint region for $D_{\bar{c}}$ is a polyhedral convex set that depends only on the nonstochastic exponent Matrix A.

Instead of working directly with D_c we shall use the transformed dual program, $\tilde{D}_c$, considered in [4, Ch. 3]. Recall $\tilde{D}_{\bar{c}}$ is obtained from $D_{\bar{c}}$ by solving the normality and orthogonality constraints of $D_{\bar{c}}$.

In what follows we shall assume without loss of generality that the rank of A is m and that $q \in R^n$ is not in the column space of A, where $q_i = 1$ if $i \leqslant n_0$ and $q_i = 0$ if $i > n_0$ (see [4, Ch. 3]). As in [4] we define d to be the dimension of the solution space of the system of equations $A'\delta = \underline{0}$, $q'\delta = 0$. (Recall d is called the degree of difficulty of $P_{\bar{c}}$ and, under the above

assumptions, equals $n - m - 1$.) Throughout the paper we assume $d > 0$. (The distribution problem for $v(P_c)$ when $d = 0$ has been studied by R. Stark in [13].) In accordance with the terminology in [4], we define $N \underline{\Delta} \{b^{(j)} \mid j \in <d>\}$ to be a nullity set for P_c iff N is a basis for the solution space of the above homogeneous system of equations. Also $b^{(0)} \in R^n$ is called a normality vector for P_c iff $A'b^{(0)} = \underline{0}$ and $q'b^{(0)} = 1$.

Let $N \underline{\Delta} \{b^{(j)} \mid j \in <d>\}$ be any nullity set and $b^{(0)}$ be any normality vector for P_c. Note $\delta \in R^n$ satisfies the orthogonality and normality conditions for D_c iff $\delta = b^{(0)} + \sum_{j=1}^{d} r_j b^{(j)}$ where the $r_j \in R^1$ are uniquely determined by δ. Thus, by replacing δ_i in D_c by $b_i^{(0)} + \sum_{j=1}^{d} r_j b_i^{(j)}$ we obtain the equivalent transformed dual problem $\tilde{D}_c$:

$$\sup_r K(\bar{c},b^{(0)}) \prod_{j=1}^{d} K(\bar{c},b^{(j)}) \prod_{i=1}^{n} \delta_i(r)^{-\delta_i(r)} \prod_{\kappa=1}^{p} \lambda_\kappa(r)^{\lambda_\kappa(r)}$$

subject to the positivity constraint $Br \geqq -b^{(0)}$ where $r \underline{\Delta} (r_1, \ldots, r_d)'$ (the vector of basic variables). In the above $\{K(\bar{c},b^{(j)}) \mid j \in <\bar{d}>\}$ is called a set of basic constants for P_c (corresponding to N and $b^{(0)}$) where $K(\bar{c},b^{(j)}) \underline{\Delta} \prod_{i=1}^{n} \bar{c}_i^{b_i^{(j)}}$ for $j \in <\bar{d}>$. Also, B is the n by d matrix whose jth column is $b^{(j)}$ for $j \in <d>$. Finally, for $i \in <n>$ and $\kappa \in <p>$, $\delta_i(r) \underline{\Delta} b_i^{(0)} + \sum_{j=1}^{d} r_j b_i^{(j)}$ and $\lambda_\kappa(r) \underline{\Delta} \sum_{i \in I_\kappa} \delta_i(r)$. When P_c has no forced constraints we define $\prod_{\kappa=1}^{p} \lambda_\kappa(r)^{\lambda_\kappa(r)}$ to be 1.

Note that the parameters in $\tilde{D}_c$ depend on the choice of nullity set N and normality vector $b^{(0)}$. However, as $v(\tilde{D}_c) = v(D_c)$ for $\bar{c} \in R^n_>$ (where $v(\tilde{D}_c)$ denotes the optimal value of $\tilde{D}_c$), the optimal value of $\tilde{D}_c$ is independent of the choice of N and $b^{(0)}$. Thus, for any nullity set N and normality vector $b^{(0)}$ for P_c, the distribution function of $v(\tilde{D}_c)$ is identical to the distribution function of $v(D_c)$, where $v(\tilde{D}_c)$ is the stochastic variable that takes on the value $v(\tilde{D}_c)$ when c_I takes on the value $\bar{c}_I$.

To obtain the distribution function and/or moments of $v(\tilde{D}_c)$ we shall first obtain the density function of the random vector $\tilde{L} \underline{\Delta} (L_0, L_1, \ldots, L_d)'$ and $L \underline{\Delta} (L_1, \ldots, L_d)'$ where, for $j \in <\bar{d}>$, L_j is the random variable that takes on the value $\log K(\bar{c},b^{(j)})$ when c_I takes on the value $\bar{c}_I$.

3. On the Density Functions of L and $\tilde{L}$

Unless otherwise stated, throughout the remainder of the paper we shall assume the following:

(1) $\{e_\nu \mid \nu \in <u>\}$ is a set of positive-valued random variables such that, for every $i \in <n>$, $c_i = \alpha_i \prod_{\nu=1}^{u} e_\nu^{\beta_{i\nu}}$ for some $\alpha_i \in (0, \infty)$ and $\beta_{i\nu} \in (-\infty, \infty)$, $\nu \in <u>$. Furthermore, it is assumed that $(\log e_1, \ldots, \log e_u)'$ is a nondegenerate normal random vector with mean vector $\mu = (\mu_1, \ldots, \mu_u)'$ and dispersion matrix Λ;

(2) There exists a nullity set $\{\hat{b}^{(j)} \mid j \in <d>\}$ for P_c such that $\{\hat{s}^{(j)} \mid j \in <d>\}$ is linearly independent where $\hat{s}^{(j)} \underline{\Delta} \beta' \hat{b}^{(j)}$ for $j \in <d>$ and β is the $n \times u$ matrix whose (i, j) entry is β_{ij};

(3) For every value $\bar{c}_l$ that c_l takes on the program P_c is superconsistent and soluble;

(4) For every value $\bar{c}_l$ that c_l takes on the program $D_{\bar{c}}$ has a unique optimal point $\delta_{\bar{c}}$ and $\delta_{\bar{c}} > 0$.

Many of the results obtained under the above restrictions form the basis to approaches for calculating the distribution function and/or moments of $v(P_c)$ under less restrictive assumptions. This will be briefly indicated in Section 6.

Assumption (1) allows for the possibility that a cost coefficient c_i is constant ($\beta_{i\nu} = 0$ for all $\nu \in <u>$). Also, (1) permits one to conveniently work with a vector of stochastic cost coefficients $c_l = (c_{i_1}, \ldots, c_{i_u})'$ for which $(\log c_{i_1}, \ldots, \log c_{i_u})'$ is a degenerate normal random vector. Degeneracy would occur, e.g., if c_1 and c_2 where components of c_l such that $c_2 = \alpha c_1^{\beta}$ for some $\alpha \in (0, \infty)$ and $\beta \in R^1$.

To evaluate the mean μ_i and variance σ_i^2 of $\log e_i$ a cost analyst could apply steps (1) through (4) of Section 1 to e_i in place of c_i. After choosing ξ_i, the median of e_i, and the variance of e_i by these steps the values of μ_i and σ_i^2 can easily be calculated [2].

Note Assumption (2) is satisfied if $u = n$ and $c_i = e_i$ for every $i \in <u>$. Also, if there exists a nullity set of P_c that satisfies (2) then every nullity set of P_c satisfies (2) (Proposition 1).

Recall, for $\bar{c} \in R^n_>$, P_c is called superconsistent iff there exists $t \in R^m$ such that $t > \underline{0}$ and $\sum_{i \in J_\kappa} \bar{c}_i \prod_{j=1}^{m} t_j^{a_{ij}} < 1$ for every $\kappa \in <p>$. Also, P_c is called soluble iff P_c has an optimal point. It can easily be shown that P_c is superconsistent for all $\bar{c} \in R^n_>$ iff there exists a linear combination of the columns of A, say x, such that $x_i < 0$ for all $i \in J_\kappa$, $\kappa \in <p>$ [1, p. 329]. Alternately, one can show that P_c is superconsistent for all $\bar{c} \in R^n_>$ iff the set $\{\delta \in R^n | \delta_i \geqslant 0 \ \forall \ i \in <n>, A'\delta = 0, \text{ and } q'\delta = 1\}$ is bounded [1, p. 329]. Moreover, if the above set is bounded and contains a point $\delta > \underline{0}$ then $P_{\bar{c}}$ will be superconsistent and soluble for every $\bar{c} \in R^n_>$ (by [4, p. 120, Th. 2] and [1, p. 329]).

Assumption (3) implies that $v(P_c) = v(D_c)$ for every value $\bar{c}_l$ taken on by c_l ([4, p. 117, Th. 1]).

Assumption (4) holds for $\bar{c} \in R^n_>$ if P_c is soluble and has no forced constraints. More generally, one can show (4) holds at $\bar{c} \in R^n_>$ if P_c is a superconsistent soluble program whose forced constraints are nonredundant and whose forced constraint gradients are linearly independent at optimality. By nonredundant we mean that the optimal value of P_c is greater than the optimal value of $P_{\kappa,c}$ for every $\kappa \in <p>$, where $P_{\kappa,c}$ denotes the program obtained from P_c by deleting forced constraint κ.

If the components of L form a set of independent random variables then we obtain a simpler formula for the density function of L since, in this case, $g(l_1, \ldots, l_d) = \prod_{i=1}^{d} g_i(l_i)$ where g is the density function for L and g_i is the density function for L_i. If, in addition, L_0 is independent of the components of L then the calculation of moments of $v(P_c)$ is simplified. This follows from the fact that one can express $v(P_c)$ as the product $e^{L_0}\omega(r)$ where ω is a known function of a d-dimensional random vector r whose density function can be calculated

from that of L. Thus, when L_0 is independent of L we have $E^{\nu}(v(P_c)) = [E^{\nu}(e^{L_0})]\,[E^{\nu}(\omega(r))]$ where $E^{\nu}(Q)$ denotes the νth moment of random variable Q (whenever this moment exists). If L_0 is a linear function of the components of L one can obtain a function $\tilde{\omega}$ of r such that $v(P_c) = \tilde{\omega}(r)$ from which one can calculate $E^{\nu}(v(P_c))$. It will be shown, under the previously listed assumptions, that one can always find a nullity set $\{b^{(j)} | j \in <d>\}$ and normality vector $b^{(0)}$ for P_c such that L_0 is independent of L if $\hat{s}^{(0)} \notin \text{span}\,\{\hat{s}^{(j)} | j \in <d>\}$ and L_0 is a linear function of the components of L if $\hat{s}^{(0)} \in \text{span}\,\{\hat{s}^{(j)} | j \in <d>\}$ where $\hat{s}^{(0)} \triangleq \beta' \hat{b}^{(0)}$ and $\hat{b}^{(0)}$ is any normality vector of P_c.

Theorem 1 indicates how to obtain a nullity set for P_c, $\{b^{(j)} | j \in <d>\}$, such that the components of the corresponding random vector L are independent normal variates whose means and variances are known. Also, using the above nullity set, it is shown how to obtain a normality vector for P_c, $b^{(0)}$, such that if $\hat{s}^{(0)} \notin \text{span}\,\{\hat{s}^{(j)} | j \in <d>\}$ then the components of the corresponding random vector $\tilde{L}$ are independent normal variates whose means and variances are known. The proof of Theorem 1 follows immediately from Lemma A which is stated and derived in Appendix A. The proof of Lemma A uses the eigenvectors of the dispersion matrix Λ. Fortunately, however, the calculation of the above-mentioned nullity set and normality vector and the calculation of the means and variances of the corresponding variates L_j, $j \in <\bar{d}>$, do not require any eigenvector or eigenvalue calculations.

THEOREM 1: (a) Define $\{b^{(j)} | j \in <d>\}$ inductively by $b^{(1)} \triangleq \hat{b}^{(1)}$ and, for $1 < j \leqslant d$, $b^{(j)} \triangleq \hat{b}^{(j)} - \sum_{l=1}^{j-1} (<\beta' b^{(l)},\ \beta' b^{(l)}>_{\Lambda})^{-1}\ (<\beta' \hat{b}^{(j)},\ \beta' b^{(l)}>_{\Lambda}) b^{(l)}$, where $<x,y>_{\Lambda} \triangleq x'\Lambda y$ for $(x,y) \in R^u \times R^u$. Then $\{b^{(j)} | j \in <d>\}$ is a well-defined nullity set of P_c.

(b) Define $b^{(0)} \triangleq \hat{b}^{(0)} - \sum_{l=1}^{d} (<\beta' b^{(l)},\ \beta' b^{(l)}>_{\Lambda})^{-1}\ (<\beta' \hat{b}^{(j)},\ \beta' b^{(l)}>_{\Lambda}) b^{(l)}$ if $\hat{s}^{(0)} \notin \text{span}\,\{\hat{s}^{(j)} | j \in <d>\}$, $b^{(0)} \triangleq \hat{b}^{(0)}$ otherwise. Then $b^{(0)}$ is a well-defined normality vector of P_c.

(c) For every $j \in <\bar{d}>$ let L_j denote the random variable that takes on the value $\log K_j(\bar{c}, b^{(j)})$ when c_l takes on the value $\bar{c}_l$. Also, define $L \triangleq (L_1, \ldots, L_d)'$ and $\tilde{L} \triangleq (L_0, L_1, \ldots, L_d)'$. Then L is a normal random vector with independent components. Additionally, $\tilde{L}$ is a normal random vector with independent components if $\hat{s}^{(0)} \notin \text{span}\,\{\hat{s}^{(j)} | j \in <d>\}$.

(d) For every $j \in <\bar{d}>$ let g_j denote the density function of L_j. Then $g_j(l) = (\eta_j \sqrt{2\pi})^{-1} \exp(-(2\eta_j^2)^{-1}(l - \nu_j)^2)$ for every $l \in R^1$, where ν_j is the expected value of L_j and η_j^2 is the variance of L_j. Furthermore, $\nu_j = <\mu, \beta' b^{(j)}> - \sum_{i=1}^{n} \alpha_i b_i^{(j)}$ and $\eta_j^2 = <\beta' b^{(j)},\ \beta' b^{(j)}>_{\Lambda}$ for every $j \in <\bar{d}>$, where $<\cdot, \cdot>$ denotes the usual inner product on R^u.

Throughout the remainder of the paper we define $b^{(j)}$ and L_j for $j \in <\bar{d}>$, L, and $\tilde{L}$ as in Theorem 1. We also denote $K_j(c, b^{(j)})$ by $K_j(c)$ for $j \in <\bar{d}>$.

We shall now show that if there exists a nullity set of P_c that satisfies Assumption (2) then every nullity set of P_c must satisfy (2).

PROPOSITION 1: If Assumption (2) holds then for any nullity set $\{\tilde{b}^{(j)} | j \in <d>\}$ of P_c the set $\{\tilde{s}^{(j)} | j \in <d>\}$ is linearly independent where $\tilde{s}^{(j)} \triangleq \beta' \tilde{b}^{(j)}$ for every $j \in <d>$.

PROOF: Let $\{\hat{b}^{(j)} | j \in <d>\}$ be a nullity set for P_c such that $\{\hat{s}^{(j)} | j \in <d>\}$ is linearly independent where $\hat{s}^{(j)} \underline{\Delta} \beta' \hat{b}^{(j)}$ for every $j \in <d>$. Let $B \underline{\Delta}$ span $\{\hat{b}^{(j)} | j \in <d>\}$ and $S \underline{\Delta}$ span $\{\hat{s}^{(j)} | j \in <d>\}$. Define T to be the unique linear transformation for B to S such that $T(\hat{b}^{(j)}) = \hat{s}^{(j)}$ for every $j \in <d>$. Thus, since $\beta' \hat{b}^{(j)} = \hat{s}^{(j)}$ for every $j \in <d>$, one has $T(b) = \beta' b$ for all $b \in B$.

Since $\{\tilde{b}^{(j)} | j \in <d>\}$ is a nullity set for P_c one has span $\{\tilde{b}^{(j)} | j \in <d>\} = B$. Thus, for every $j \in <d>$, $T(\tilde{b}^{(j)}) = \beta' \tilde{b}^{(j)} = \tilde{s}^{(j)}$. Note T is an isomorphism from B onto S and $\{\tilde{b}^{(j)} | j \in <d>\}$ is linearly independent. Hence, $\{\tilde{s}^{(j)} | j \in <d>\}$ is linearly independent.

Next we consider the assumption $\hat{s}^{(0)} \notin$ span $\{\hat{s}^{(j)} | j \in <d>\}$.

PROPOSITION 2: Assume $\hat{s}^{(0)} \notin$ span $\{\hat{s}^{(j)} | j \in <d>\}$. Then for any nullity set $\{\tilde{b}^{(j)} | j \in <d>\}$ and normality vector $\tilde{b}^{(0)}$ for P_c one has $\tilde{s}^{(0)} \notin$ span $\{\tilde{s}(j) | j \in <d>\}$ where $\tilde{s}^{(j)} \underline{\Delta} \beta' \tilde{b}^{(j)}$ for every $j \in <\bar{d}>$.

PROOF: Define $\hat{B} \underline{\Delta}$ span $\{\hat{b}^{(j)} | j \in <\bar{d}>\}$ and $\hat{S} \underline{\Delta}$ span $\{\hat{s}^{(j)} | j \in <\bar{d}>\}$. Let $\hat{T}$ be the unique linear transformation from B to S for which $\hat{T}(\hat{b}^{(j)}) = \hat{s}^{(j)}$ for every $j \in <\bar{d}>$. Since $\beta' \hat{b}^{(j)} = \hat{s}^{(j)}$ for every $j \in <\bar{d}>$ one has $\hat{T}(b) = \beta' b$ for all $b \in \hat{B}$.

Since $\{\tilde{b}^{(j)} | j \in <d>\}$ is a nullity set of P_c one has span $\{\tilde{b}^{(j)} | j \in <d>\} \subset \hat{B}$. Also, since $\tilde{b}^{(0)}$ and $\hat{b}^{(0)}$ are normality vectors of P_c it follows that $\tilde{b}^{(0)} - \hat{b}^{(0)} \in$ span $\{\hat{b}^{(j)} | j \in <d>\}$, i.e., $\tilde{b}^{(0)} \in \hat{B}$. Thus, for every $j \in <\bar{d}>$, $\tilde{b}^{(j)} \in \hat{B}$ and hence, $\hat{T}(\tilde{b}^{(j)}) = \beta' \tilde{b}^{(j)} = \tilde{s}^{(j)}$. Finally, observe $\hat{T}$ is an isomorphism from $\hat{B}$ onto $\hat{S}$ since $\hat{s}^{(0)} \notin$ span $\{\hat{s}^{(j)} | j \in <d>\}$ and $\{\hat{s}^{(j)} | j \in <d>\}$ is linearly independent by Assumption (2). Moreover, $\{\tilde{b}^{(j)} | j \in <\bar{d}>\}$ is linearly independent. Thus $\{\tilde{s}^{(j)} | j \in <\bar{d}>\}$ is linearly independent.

As mentioned earlier, when $u = n$ and $c_i = e_i$ for every $i \in <u>$ then Assumption (2) holds. In addition one has $\hat{s}^{(0)} \notin$ span $\{\hat{s}^{(j)} | j \in <d>\}$ and hence by Theorem 1 the components of $\tilde{L}$ are independent. We next consider the case where Assumption (2) holds but $\hat{s}^{(0)} \in$ span $\{\hat{s}^{(j)} | j \in <d>\}$.

PROPOSITION 3: Assume $\hat{s}^{(0)} \in$ span $\{\hat{s}^{(j)} | j \in <d>\}$. Then there exist $y_j \in R^1$ for $j \in <d>$ such that $s^{(0)} = \sum_{j=1}^{d} y_j s^{(j)}$ where $s^{(j)} \underline{\Delta} \beta' b^{(j)}$ for every $j \in <\bar{d}>$. Furthermore, $L_0 = \sum_{j=1}^{d} y_j L_j + D$ where D is the constant $\sum_{i=1}^{n} \left[b_i^{(0)} - \sum_{j=1}^{d} y_j b_i^{(j)} \right] \log \alpha_i$.

PROOF: Since $\hat{s}^{(0)} \in$ span $\{\hat{s}^{(j)} | j \in <d>\}$, by Proposition 2 there exist $y_j \in R^1$ for $j \in <d>$ such that $s^{(0)} = \sum_{j=1}^{d} y_j s^{(j)}$. Thus $s_i^{(0)} = \sum_{j=1}^{d} y_j s_i^{(j)}$ for every $i \in <u>$.

By Lemma A, Part (iii), $L_0 \underline{\Delta} \log K_0(c) = \log \left[\left(\prod_{i=1}^{n} \alpha_i^{b_i^{(0)}} \right) \{\exp(L(e, s^{(0)}))\} \right] = \sum_{i=1}^{n} b_i^{(0)} \log \alpha_i + \sum_{i=1}^{u} s_i^{(0)} \log e_i = \sum_{i=1}^{n} b_i^{(0)} \log \alpha_i + \sum_{i=1}^{u} \left[\sum_{j=1}^{d} y_j s_i^{(j)} \right] \log e_i = \sum_{i=1}^{n} b_i^{(0)} \log \alpha_i + \sum_{j=1}^{d} \left(y_j \sum_{i=1}^{u} s_i^{(j)} \log e_i \right) = \sum_{i=1}^{n} b_i^{(0)} \log \alpha_i + \sum_{j=1}^{d} y_j L(e, s^{(j)}) = \sum_{i=1}^{n} b_i^{(0)} \log \alpha_i + \sum_{j=1}^{d} y_j$

$\left(L_j - \sum_{i=1}^{n} b_i^{(j)} \log \alpha_i\right) = \sum_{j=1}^{d} y_j L_j + D$ where $D \triangleq \sum_{i=1}^{n} b_i^{(0)} \log \alpha_i - \sum_{j=1}^{d} y_j \left(\sum_{i=1}^{n} b_i^{(j)} \log \alpha_i\right) = \sum_{i=1}^{n} \left(b_i^{(0)} - \sum_{j=1}^{d} y_j b_i^{(j)}\right) \log \alpha_i$.

We next consider the d-dimensional random vector $r \triangleq (r_1, \ldots, r_d)'$ mentioned earlier whose density function can be used to obtain moments of $v(P_c)$. To define r assume c_l takes on the value $\bar{c}_l$. Then, by Assumption (4), $D_{\bar{c}}$ has a unique optimal point $\delta_{\bar{c}}$. Since $\{b^{(j)} | j \in <d>\}$ is a nullity set and $b^{(0)}$ is a normality vector of $P_{\bar{c}}$ there exists a unique point $r_{\bar{c}} \triangleq (r_1(\bar{c}), \ldots, r_d(\bar{c}))' \in R^d$ for which $\delta_{\bar{c}} = b^{(0)} + \sum_{j=1}^{d} (r_j(\bar{c})) b^{(j)}$. We define r to be the random vector that takes on the value $r_{\bar{c}}$ when c_l takes on the value $\bar{c}_l$. In the next section we shall obtain the density function of r from that of L. Also, when $\hat{s}^{(0)} \notin \text{span } \{\hat{s}^{(j)} | j \in <d>\}$, we shall obtain the density function of $v(P_c)$ as a marginal density of $(v(P_c), r)$. The density function of $(v(P_c), r)$ is obtained from that of $\tilde{L}$.

4. THE DENSITY FUNCTIONS OF r and $v(P_c)$

Assume c_l takes on the value $\bar{c}_l$. Since $\delta_{\bar{c}}$ is an optimal point for $D_{\bar{c}}$ with all positive components (Assumption (4)) it follows that $\delta_{\bar{c}}$ satisfies the maximizing equations for $D_{\bar{c}}$ [4, p. 88, Th. 3]. Expressing $\delta_{\bar{c}}$ in terms of the components of $r_{\bar{c}}$, the maximizing equations can be written in the form $\log K_j(\bar{c}) = h_j(r_{\bar{c}})$ for every $j \in <d>$ where, for $j \in <d>$, h_j is the function defined in Theorem 2. The above equations will be used to obtain the density function of r from that of L. From the above maximizing equations one can easily show that the optimal value of $P_{\bar{c}}$ satisfies the equation $\log K_0(\bar{c}) = \log (v(P_{\bar{c}})) + h_0(r_{\bar{c}})$ where h_0 is defined as in Theorem 2 [4, p. 88, Th. 3]. This equation, together with the maximizing equations, will be used to obtain the density function of $(v(P_c), r)$ from that of $\tilde{L}$ when $\hat{s}^{(0)} \notin \text{span } \{\hat{s}^{(j)} | j \in <d>\}$.

To obtain the density functions of r and $(v(P_c), r)$ we shall first define the functions h_j for $j \in <\bar{d}>$ and establish several of their properties.

THEOREM 2: Let $H \triangleq \{r \in R^d | Br > -b^{(0)}\}$ where B is the $n \times d$ matrix whose jth column is $b^{(j)}$ for $j \in <d>$. For every $j \in <\bar{d}>$ define $h_j: H \to R^1$ such that, for $r \in H$,

$$h_j(r) \triangleq \sum_{i=1}^{n} b_i^{(j)} \log \delta_i(r) - \sum_{\kappa=1}^{p} \lambda_\kappa^{(j)} \log \lambda_\kappa(r)$$

(where $\sum_{\kappa=1}^{p} \lambda_\kappa^{(j)} \log \lambda_\kappa(r) \triangleq 0$ if $p = 0$). In the above, for every $r \in R^d$, $\delta_i(r) \triangleq b_i^{(0)} + \sum_{j=1}^{d} r_j b_i^{(j)}$ for $i \in <n>$ and $\lambda_\kappa(r) \triangleq \sum_{i \in J_\kappa} \delta_i(r)$ for $\kappa \in <p>$. Also, for every $j \in <\bar{d}>$ and $\kappa \in <\bar{p}>$, $\lambda_\kappa^{(j)} \triangleq \sum_{i \in J_\kappa} b_i^{(j)}$.

For every $j \in <\bar{d}>$ define $\tilde{h}_j: (0, \infty) \times H \to R^{d+1}$ such that, for $(z, r) \in (0, \infty) \times H$, $\tilde{h}_j(z, r) \triangleq h_j(r)$ if $j \in <d>$ and $\tilde{h}_0(z, r) \triangleq \log z + h_0(r)$.

Finally, define $h: H \to R^d$ and $\tilde{h}: (0, \infty) \times H \to R^{d+1}$ such that, for every $(z, r) \in (0, \infty) \times H$, $h(r) \triangleq (h_1(r), \ldots, h_d(r))'$ and $\tilde{h}(z, r) \triangleq (\tilde{h}_0(z, r), \tilde{h}_1(z, r), \ldots, \tilde{h}_d(z, r))'$ Then

(a) h and $\tilde{h}$ are continuously differentiable in H and $(0,\infty) \times H$ respectively;

(b) h is 1-1 in H and $\tilde{h}$ is 1-1 in $(0,\infty) \times H$;

(c) h is onto R^d;

(d) If $\hat{s}^{(0)} \notin \text{span}\{\hat{s}^{(j)} | j \in <d>\}$ then $\tilde{h}$ is onto R^{d+1}.

PROOF: (a) Clearly, for every $j \in <d>$, h_j is continuously differentiable in H and, for every $j \in <\bar{d}>$, $\tilde{h}_j$ is continuously differentiable in $(0,\infty) \times H$. Thus, h is continuously differentiable in H and $\tilde{h}$ is continuously differentiable in $(0,\infty) \times H$.

(b) Let r and s be elements of H such that $h(r) = h(s)$. Note $\delta(r) > \underline{0}$ and $\delta(s) > \underline{0}$. Also, since $(\log K_1(c), \ldots, \log K_d(c))'$ is a nondegenerate d-dimensional normal random vector, c_j takes on a value, say $\bar{c}_j$, for which $\log K_j(\bar{c}) = h_j(r) = h_j(s)$ for every $j \in <d>$. Thus, by definition of K_j and h_j for $j \in <d>$, $\delta(r)$ and $\delta(s)$ satisfy the maximizing equations for $D_{\bar{c}}$. Also, $\delta(r)$ and $\delta(s)$ are feasible points of $D_{\bar{c}}$. Thus, by [4, p. 88, Th. 3], $\delta(r)$ and $\delta(s)$ are optimal points of $D_{\bar{c}}$. However, by Assumption (4), $D_{\bar{c}}$ has only one optimal point and hence $\delta(r) = \delta(s)$. This implies $r = s$ since the nullity set $\{b^{(j)} | j \in <d>\}$ is linearly independent.

Next, let (z_1,r) and (z_2,s) be elements of $(0,\infty) \times H$ such that $\tilde{h}(z_1,r) = \tilde{h}(z_2,s)$. Then, $h(r) = h(s)$, and hence, $r = s$. Also, $\tilde{h}_0(z_1,r) = \tilde{h}_0(z_2,s)$. Hence, by the definition of $\tilde{h}_0$,

$$\begin{aligned}\log z_1 &= \tilde{h}_0(z_1,r) - \sum_{i=1}^{n} b_i^{(0)} \log \delta_i(r) + \sum_{\kappa=1}^{p} \lambda_\kappa^{(0)} \log \lambda_\kappa(r) \\ &= \tilde{h}_0(z_2,s) - \sum_{i=1}^{n} b_i^{(0)} \log \delta_i(s) + \sum_{\kappa=1}^{p} \lambda_\kappa^{(0)} \log \lambda_\kappa(s) \\ &= \log z_2\end{aligned}$$

and thus $z_1 = z_2$.

(c) Let $u \in R^d$. Since $(\log K_1(c), \ldots, \log K_d(c))'$ is a nondegenerate d-dimensional normal random vector, c_j takes on a value, say $\bar{c}_j$, for which $\log K_j(\bar{c}) = u_j$ for every $j \in <d>$. By Assumption (4), $D_{\bar{c}}$ has an optimal point $\hat{\delta}$ such that $\hat{\delta} > \underline{0}$. Let r be the unique element of R^d for which $\hat{\delta} = b^{(0)} + \sum_{i=1}^{d} r_i b^{(i)}$ and denote $\hat{\delta}$ by $\delta(r)$.

Since $\delta(r) > \underline{0}$ one has $r \in H$. Furthermore, since $\delta(r)$ is an optimal point of $D_{\bar{c}}$, by [4, p. 88, Th. 3] $\delta(r)$ satisfies the maximizing equations for $D_{\bar{c}}$. Thus, for every $j \in <d>$, $h_j(r) = \log K_j(\bar{c}) = u_j$. Hence, h is onto R^d.

(d) Assume $\hat{s}^{(0)} \notin \text{span}\{\hat{s}^{(j)} | j \in <d>\}$. Let $u = (u_1, \ldots, u_d)' \in R^d$ and $u_0 \in R^1$. By Theorem 1, $(K_0(c), K_1(c), \ldots, K_d(c))'$ is a nondegenerate $(d+1)$-dimensional normal random vector. Thus, c_j takes on a value, say $\bar{c}_j$, for which $\log K_j(\bar{c}) = u_j$ for every $j \in <\bar{d}>$. By Assumption (4), $D_{\bar{c}}$ has an optimal point $\hat{\delta}$ such that $\hat{\delta} > \underline{0}$. Let r be the unique element of R^d for which $\hat{\delta} = b^{(0)} + \sum_{i=1}^{d} r_i b^{(i)}$ and denote $\hat{\delta}$ by $\delta(r)$. Let $r_0 \triangleq v(D_{\bar{c}})$.

Since $\delta(r) > \underline{0}$ and $v(D_{\bar{c}}) > 0$ one has $(r_0, r) \in (0,\infty) \times H$. Also, since $\delta(r)$ is an optimal point of $D_{\bar{c}}$, by [4, p. 88, Th. 3] $\delta(r)$ satisfies the maximizing equations for $D_{\bar{c}}$. Thus, for every $j \in <d>$,

$$\tilde{h}_j(r_0, r) = h_j(r) = \log K_j(\bar{c}) = u_j. \tag{1}$$

Also, by [4, p. 88, Th. 3], since $\delta(r)$ satisfies the maximizing equations for D_c one has

$$r_0 = v(D_{\bar{c}}) = K_0(\bar{c}) \prod_{i=1}^{n} \delta_i(r)^{-b_i^{(0)}} \prod_{\kappa=1}^{p} \lambda_\kappa(r)^{\lambda_\kappa^{(0)}}.$$

Thus, $\log r_0 = \log K_0(\bar{c}) - \sum_{i=1}^{n} b_i^{(0)} \log \delta_i(r) + \sum_{\kappa=1}^{p} \lambda_\kappa^{(0)} \log (\lambda_\kappa(r))$, i.e.

$$\tilde{h}_0(r_0, r) = \log K_0(\bar{c}) = u_0. \tag{2}$$

By (1) and (2) $\tilde{h}$ is onto R^{d+1}.

Note by Theorem 2, for every $l \in R^d$ there exists a unique point $r_l \in H$ such that $l = h(r_l)$. Thus, we can define $h^{-1}: R^d \to H$ by $h^{-1}(l) \underline{\Delta} r_l$ for $l \in R^d$. Also, if $\hat{s}^{(0)} \notin$ span $\{\hat{s}^{(j)} | j \in <d>\}$ then by Theorem 2, for every $\tilde{l} \in R^{d+1}$ there exists a unique point $(z_{\tilde{l}}, r_{\tilde{l}}) \in (0,\infty) \times H$ such that $\tilde{l} = \tilde{h}(z_{\tilde{l}}, r_{\tilde{l}})$. Thus, when $\hat{s}^{(0)} \notin$ span $\{\hat{s}^{(j)} | j \in <d>\}$, we can define $\tilde{h}^{-1}: R^{d+1} \to (0,\infty) \times H$ by $\tilde{h}^{-1}(\tilde{l}) \underline{\Delta} (z_{\tilde{l}}, r_{\tilde{l}})$ for $\tilde{l} \in R^{d+1}$.

PROPOSITION 4: (a) $r = h^{-1}(L)$;

(b) $(v(P_c), r) = \tilde{h}^{-1}(\tilde{L})$ if $\hat{s}^{(0)} \notin$ span $\{\hat{s}^{(j)} | j \in <d>\}$.

PROOF: (a) Let c_l take on the value $\bar{c}_l$. Then L takes on the value $l \underline{\Delta} (\log K_1(\bar{c}), \ldots, \log K_d(\bar{c}))'$. By Assumption (4) and [4, p. 88, Th. 3] one has $r_{\bar{c}} \in H$ and $\log K_j(\bar{c}) = h_j(r_{\bar{c}})$ for every $j \in <d>$. Thus, $h^{-1}(l) = r_{\bar{c}}$. Hence, $h^{-1}(L)$ takes on the value $r_{\bar{c}}$ when c_l takes on the value $\bar{c}_l$, i.e., $h^{-1}(L) = r$.

(b) Assume $\hat{s}^{(0)} \notin$ span $\{\hat{s}^{(j)} | j \in <d>\}$. Then $\tilde{h}^{-1}: R^{d+1} \to (0,\infty) \times H$ is well-defined. Let c_l take on the value $\bar{c}_l$. Then $\tilde{L}$ takes on the value $\tilde{l} \underline{\Delta} (\log K_0(\bar{c}), \log K_1(\bar{c}), \ldots, \log K_d(\bar{c}))'$. Note $v(P_c)$ takes on the value $v(P_{\bar{c}}) > 0$. Also, by Assumption (4) and [4, p. 88, Th. 3] one has $r_{\bar{c}} \in H$ and $\log K_j(\bar{c}) = \tilde{h}_j(v(P_{\bar{c}}), r_{\bar{c}})$ for every $j \in <\bar{d}>$ Thus, $\tilde{h}^{-1}(\tilde{l}) = (v(P_{\bar{c}}), r_{\bar{c}})$. Hence, $\tilde{h}^{-1}(\tilde{L})$ takes on the value $(v(P_c), r_c)$ when c_l takes on the value $\bar{c}_l$, i.e., $\tilde{h}^{-1}(\tilde{L}) = (v(P_c), r)$.

We can now obtain the density function of r.

THEOREM 3: Let ψ denote the density function of r and g denote the density function of L. Then

$$\psi(r) = \begin{cases} 0 & \text{if } r \notin H, \\ \{g(h(r))\}(|\det Dh(r)|) & \text{if } r \in H, \end{cases}$$

where $Dh(r)$ denotes the derivative of h at r.

PROOF: Let c_j take on the value $\bar{c}_j$. Then r takes on the value $r_0 \triangleq (r_1, \ldots, r_d)'$ where $\delta_0 = b^{(0)} + \sum_1^d r_j b^{(j)}$ is the unique optimal point of D_0. By Assumption (4), $\delta_0 > \underline{0}$. Thus, $Br_0 > -b^{(0)}$, i.e., $r_0 \in H$. Hence, r can only take on values in H. Thus, $\psi(r) = 0$ if $r \notin H$.

Let B be an open Borel subset of H. Note, by Proposition 4,

$$Pr(r \in B) = Pr(h^{-1}(L) \in B) = Pr(L \in h(B)).$$

By Theorem 2, h is 1-1 and continuously differentiable in B. Also, g is integrable on $h(B)$ since g is the density function of L. Thus,

$$Pr(L \in h(B)) = \int_{h(B)} g = \int_B (g \circ h)|\det Dh|$$

[12, Ths. 3-13 and 3-14], where $g \circ h$ denotes the composition of g and h. Hence, $Pr(r \in B) = \int_B (g \circ h)|\det Dh|$. This implies $\psi(r) = \{g(h(r))\}(|\det Dh(r)|)$ for $r \in H$.

Next we obtain the density function of $(v(P_0), r)$.

THEOREM 4: Let $\tilde{\psi}$ denote the density function of $(v(P_0), r)$ and $\tilde{g}$ denote the density function of $\tilde{L}$. Assume $\hat{s}^{(0)} \notin \text{span}\{\hat{s}^{(j)} | j \in \langle d \rangle\}$. Then

$$\tilde{\psi}(z,r) = \begin{cases} 0 & \text{if } (z,r) \notin (0,\infty) \times H, \\ \{\tilde{g}(\tilde{h}(z,r))\}(|\det D\tilde{h}(z,r)|) & \text{if } (z,r) \in (0,\infty) \times H, \end{cases}$$

where $D\tilde{h}(z,r)$ denotes the derivative of $\tilde{h}$ at (z,r).

PROOF: By the proof of Theorem 3 it follows that $(v(P_0), r)$ can only take on values in $(0,\infty) \times H$. Hence, $\tilde{\psi}(z,r) = 0$ if $(z,r) \notin (0,\infty) \times H$.

Let $\tilde{B}$ be an open Borel subset of $(0,\infty) \times H$ and define $z \triangleq v(P_0)$. Note by Proposition 4, $Pr((z,r) \in \tilde{B}) = Pr(\tilde{h}^{-1}(\tilde{L}) \in \tilde{B}) = Pr(\tilde{L} \in \tilde{h}(\tilde{B}))$. By Theorem 2, $\tilde{h}$ is 1-1 and continuously differentiable in $\tilde{B}$. Also, $\tilde{g}$ is integrable on $\tilde{h}(\tilde{B})$ since $\tilde{g}$ is the density function of $\tilde{L}$. Thus, $Pr(\tilde{L} \in \tilde{h}(\tilde{B})) = \int_{\tilde{h}(\tilde{B})} \tilde{g} = \int_{\tilde{B}} (\tilde{g} \circ \tilde{h})|\det D\tilde{h}|$ [12, Ths. 3-13 and 3-14], where $\tilde{g} \circ \tilde{h}$ denotes the composition of $\tilde{g}$ and $\tilde{h}$. Hence, $Pr((z,r) \in \tilde{B}) = \int_{\tilde{B}} (\tilde{g} \circ \tilde{h})\ |\det D\tilde{h}|$. This implies $\tilde{\psi}(z,r) = \{\tilde{g}(\tilde{h}(z,r))\}(|\det D\tilde{h}(z,r)|)$ for $(z,r) \in (0,\infty) \times H$.

When $\hat{s}^{(0)} \notin \text{span}\{\hat{s}^{(j)} | j \in \langle d \rangle\}$ the above theorem immediately yields the density function of $v(P_0)$.

COROLLARY 4.1: Let f denote the density function of $v(P_0)$ and assume $\hat{s}^{(0)} \notin \text{span}\{\hat{s}^{(j)} | j \in \langle d \rangle\}$. Then

$$f(z) = \begin{cases} 0 & \text{if } z \notin (0,\infty), \\ \int_{\cdot \cdot H} \{\tilde{g}(\tilde{h}(z,r))\}(|\det D\tilde{h}(z,r)|)\,dr & \text{if } z \in (0,\infty). \end{cases}$$

Observe that to evaluate f at $z \in (0,\infty)$ by Corollary 4.1 one must integrate a specified function over the convex polyhedral set $H = \{r \in R^d | Br > -b^{(0)}\}$. When the degree of difficulty d equals 1 then H will be an interval in R^1 whose end points can easily be obtained

Thus, when $d = 1$, one can accurately approximate $f(z)$ by applying a quadrature formula to evaluate the integral expression for $f(z)$. However, the quadrature rule must be modified as in [7, Ch. 7, Sec. 6.2] to allow for the fact that the integrand is not defined at the end points of the interval of integration.

When $d > 1$ the effort and expense of devising and applying a quadrature scheme to approximate the integral expression for $f(z)$ to a high degree of accuracy may not be justified since frequently the distributions chosen for the stochastic c_i will be subjectively determined. In such cases a numerical Monte Carlo method could be an attractive alternative for approximating the multiple integral used to express $f(z)$ [6, 14, 15].

Finally, under the assumption of Corollary 4.1 note the distribution function of $v(P_c)$, denoted by F, is given by

$$F(y) = \begin{cases} 0 & \text{if } y \leqslant 0, \\ \int_{(z,r)\in(0,y)\times H} \{\tilde{g}(\tilde{h}(z,r))\}(|\det D\tilde{h}(z,r)|)\,drdz & \text{if } y > 0. \end{cases}$$

Thus, if great precision is not required a numerical Monte Carlo technique could be attractive for approximating $F(y)$ as well as $f(z)$.

5. THE MOMENTS OF $v(P_c)$

In the following, for each random variable Q, recall $E^{(\nu)}(Q)$ denotes the moment of order ν of Q whenever it exists, where $\nu \in N$ (the set of positive integers). Also, let $E(Q) \underline{\Delta} E^{(1)}(Q)$.

Throughout Section 5 we assume $E^{(\nu)}(v(P_c))$ exists for every $\nu \in N$. Proposition B in Appendix B establishes that boundedness of the dual feasible set $F \underline{\Delta} \{\delta \in R^n | A'\delta = \underline{0}, q'\delta = 1, \delta_i \geqslant 0 \ \forall\ i \in <n>\}$ is a sufficient condition for the above moments to exist. Furthermore, one can show P_c is superconsistent for every $\bar{c} \in R^n_>$ iff F is bounded (see p. 554).

To calculate the moments of $v(P_c)$ it is advantageous to use the density function of r instead of that for $v(P_c)$. To obtain the moments of $v(P_c)$ in terms of the density function of r we shall use the function ω: $H \to R^1$ defined by $\omega(r) \underline{\Delta} e^{h_0(r)}$ for $r \in H$. Thus, $\omega(r) = \prod_{i=1}^{n} \delta_i(r)^{b_i^{(0)}}$ if $p = 0$ and $\omega(r) = \prod_{i=1}^{n} \delta_i(r)^{-b_i^{(0)}} \prod_{\kappa=1}^{p} \lambda_\kappa(r)^{\lambda_\kappa^{(0)}}$ if $p > 0$, for $r \in H$.

PROPOSITION 5: $v(P_c) = e^{L_0}\omega(r)$.

PROOF: Let c_j take on the value $\bar{c}_j$. Then $e^{L_0}\omega(r)$ takes on the value $K_0(\bar{c})\omega(r_c)$ where $r_c = (r_1, \ldots, r_d)'$ is the point in H for which $\delta_c = b^{(0)} + \sum_{j=1}^{d} r_j b^{(j)}$ is the unique optimal point of D_c. Since $\delta_c > \underline{0}$ one has $K_0(\bar{c})\omega(r_c) = v(P_c)$ [4, p. 88, Th. 3]. Thus, $v(P_c) = e^{L_0}\omega(r)$.

We shall now obtain the moments of $v(P_c)$ when $\hat{s}^{(0)} \in \text{span}\{\hat{s}^{(j)} | j \in <d>\}$.

THEOREM 5: Assume $\hat{s}^{(0)} \in \text{span}\{\hat{s}^{(j)} | j \in <d>\}$. Then, for every $\nu \in N$, $E^{(\nu)}(v(P_c)) = \int_{r\in H} \{\tilde{\omega}(r)\}^\nu \psi(r)\,dr$ where, for $r \in H$,

$$\tilde{\omega}(r) \triangleq \begin{cases} e^{D} \prod_{i=1}^{n} \delta_i(r)^{u_i} & \text{if } p = 0, \\ e^{D} \prod_{i=1}^{n} \delta_i(r)^{u_i} \prod_{\kappa=1}^{p} \lambda_\kappa(r)^{-v_\kappa} & \text{if } p > 0. \end{cases}$$

In the above $D \triangleq \sum_{i=1}^{n} \left[b_i^{(0)} - \sum_{j=1}^{d} y_j b_i^{(j)} \right] \log \alpha_i$, $u_i \triangleq \sum_{j=1}^{d} y_j b_i^{(j)} - b_i^{(0)}$ for $i \in <n>$, and $v_\kappa \triangleq \sum_{j=1}^{d} y_j \lambda_\kappa^{(j)} - \lambda_\kappa^{(0)}$ for $\kappa \in <p>$ where $(y_1, \ldots, y_d)'$ is the unique element of R^d for which $s^{(0)} = \sum_{j=1}^{d} y_j s^{(j)}$.

PROOF: We shall assume $p > 0$. (The modification needed in the proof for $p = 0$ should be clear.) By Assumption (2) and Proposition 1 the set $\{s^{(j)} | j \in <d>\}$ is linearly independent. Hence, by Proposition 3 there exists a unique element $(y_1, \ldots, y_d)'$ of R^d for which $s^{(0)} = \sum_{j=1}^{d} y_j s^{(j)}$. Also, by Proposition 3, $\log K_0(c) = \sum_{j=1}^{d} y_j \log K_j(c) + D$ where $D \triangleq \sum_{i=1}^{n} \left[b_i^{(0)} - \sum_{j=1}^{d} y_j b_i^{(j)} \right] \log \alpha_i$. By Proposition 5 $v(P_c) = K_0(c)\omega(r)$. Thus,

$$\begin{aligned} (1) \qquad \log (v(P_c)) &= \log K_0(c) + \log (\omega(r)) \\ &= \sum_{j=1}^{d} y_j \log K_j(c) + D + \sum_{\kappa=1}^{p} \lambda_\kappa^{(0)} \log \lambda_\kappa(r) - \sum_{i=1}^{n} b_i^{(0)} \log \delta_i(r). \end{aligned}$$

Let c_j take on the value $\bar{c}_j$. Then r takes on the value $r_c = (r_1(\bar{c}), \ldots, r_d(\bar{c}))'$. Since $\delta_{r_c} \triangleq b^{(0)} + \sum_{j=1}^{d} (r_j(\bar{c})) b^{(j)}$ is an optimal point of D_{r_c} and $\delta_{r_c} > \underline{0}$ (by Assumption (4)) one has (by [4, p. 88, Th. 3])

$$\log K_j(\bar{c}) = h_j(r_c) = \sum_{i=1}^{n} b_i^{(j)} \log \delta_i(r_c) - \sum_{\kappa=1}^{p} \lambda_\kappa^{(j)} \log \lambda_\kappa(r_c)$$

for every $j \in <d>$. Thus, by (1),

$$\begin{aligned} \log (v(P_c)) &= \sum_{j=1}^{d} y_j \left\{ \sum_{i=1}^{n} b_i^{(j)} \log \delta_i(r_c) - \sum_{\kappa=1}^{p} \lambda_\kappa^{(j)} \log \lambda_\kappa(r_c) \right\} \\ &\quad + D + \sum_{\kappa=1}^{p} \lambda_\kappa^{(0)} \log \lambda_\kappa(r_c) - \sum_{i=1}^{n} b_i^{(0)} \log \delta_i(r_c) \\ &= D + \sum_{i=1}^{n} \left[\sum_{j=1}^{d} y_j b_i^{(j)} - b_i^{(0)} \right] \log \delta_i(r_c) \\ &\quad - \sum_{\kappa=1}^{p} \left[\sum_{j=1}^{d} y_j \lambda_\kappa^{(j)} - \lambda_\kappa^{(0)} \right] \log \lambda_\kappa(r_c). \end{aligned}$$

Hence,

$$v(P_c) = e^{D} \prod_{i=1}^{n} \delta_i(r)^{u_i} \prod_{\kappa=1}^{p} \lambda_\kappa(r)^{-v_\kappa} = \tilde{\omega}(r).$$

It follows that $E^{(r)}(v(P_c)) = E(\{v(P_c)\}^r) = \int_{rH} \{\tilde{\omega}(r)\}^r \psi(r) dr$ [8, p. 18, Th. 1.4.3]

We next obtain the moments of $v(P_c)$ in terms of the density function of r when $\hat{s}^{(0)} \notin$ span $\{\hat{s}^{(j)} | j \in <d>\}$.

THEOREM 6: Assume $\hat{s}^{(0)} \notin$ span $\{\hat{s}^{(j)} | j \in <d>\}$ and let $\nu \in N$. Then,

(a) e^{L_0} is lognormal and independent of $R \triangleq \omega(r)$;

(b) $E^{(\nu)}(v(P_c)) = E^{(\nu)}(e^{L_0}) E^{(\nu)}(R)$;

(c) $E^{(\nu)}(e^{L_0}) = \exp\left[\nu\left(\sum_{i=1}^{n} \mu_i s_i^{(0)} - \sum_{i=1}^{n} \alpha_i b_i^{(0)}\right) + \frac{\nu^2}{2}(s^{(0)} \Lambda s^{(0)})\right]$

where $s^{(0)} = \beta' b^{(0)}$;

(d) $E^{(\nu)}(R) = \int_{r \in H} \{\omega(r)\}^{\nu} \psi(r)\, dr$.

PROOF: (a) By Theorem 1 L_0 is normal and hence e^{L_0} is lognormal.

Note by Proposition 4 $\omega(r) = \omega(h^{-1}(L))$ where $L = (L_1, \ldots, L_d)'$. By Theorem 1 L_0 and L are independent. Thus, e^{L_0} and R are independent [8, p. 15, (III)].

(b) To show that $E(R^{\nu})$ exists let $X \triangleq e^{\nu L_0}$ and $Y \triangleq R^{\nu}$. Clearly (X, Y) is a continuous random vector. Thus, let w be the joint density function of (X, Y). Also, let w_1 and w_2 denote the marginal densities of X and Y respectively. Then $w(x,y) = w_1(x) w_2(y)$ for all $(x,y) \in R^1 \times R^1$ since X and Y are independent by (a).

By Proposition 5 $\{v(P_c)\}^{\nu} = XY$. Thus, by assumption, $E(XY)$ exists and hence, $\iint_{(x,y) \in R^1 \times R^1} xyw(x,y)\, dx dy$ is convergent. Thus, by Fubini's Theorem [10, p. 207, Th. 2.8.7] $\iint_{(x,y) \in R^1 \times R^1} xyw(x,y)\, dx dy = \int_0^{\infty} H(y)\, dy$ where $H(y) \triangleq \int_0^{\infty} xyw(x,y)\, dx = yw_2(y) \int_0^{\infty} xw_1(x)\, dx = yw_2(y) E(X)$. This implies $E(XY) = E(X) \int_0^{\infty} yw_2(y)\, dy$ and hence $\int_0^{\infty} yw_2(y)\, dy$ is convergent, i.e., $E(R^{\nu})$ exists.

Since the expected values of $e^{\nu L_0}$ and R^{ν} exist, the independence of $e^{\nu L_0}$ and R^{ν} implies $E(e^{\nu L_0}) E(R^{\nu}) = E(e^{\nu L_0} R^{\nu})$ [5, p. 82, Th. 3.6.2]. Thus, by Proposition 5, $E^{(\nu)}(v(P_c)) = E(\{e^{L_0} R\}^{\nu}) = E(e^{\nu L_0}) E(R^{\nu}) = E^{(\nu)}(e^{L_0}) E^{(\nu)}(R)$.

(c) Recall by Theorem 1 $E(L_0) = \sum_{i=1}^{n} \mu_i s_i^{(0)} - \sum_{i=1}^{n} \alpha_i b_i^{(0)}$ and $V(L_0) = s^{(0)} \Lambda s^{(0)}$ where $s^{(0)} = \beta' b^{(0)}$ and $V(L_0)$ denotes the variance of L_0. By [2] one has $E^{(\nu)}(e^{\nu L_0}) = \exp[\nu E(L_0) + \frac{1}{2} \nu^2 V(L_0)]$ since L_0 is normal.

(d) Using the density function ψ of r we obtain $E^{(\nu)}(R) = E(R^{\nu}) = \int_{r \in H} \{\omega(r)\}^{\nu} \psi(r)\, dr$ [8, p. 18, Th. 1.4.3].

Note that to evaluate $E^{(\nu)}(v(P_c))$ by Theorem 5 or 6 one must integrate a specified function over the convex polyhedral set $H = \{r \in R^d | Br > -b^{(0)}\}$. Hence, the comments made in Section 4 concerning the evaluation of $f(z)$ also apply to the evaluation of $E^{(\nu)}(v(P_c))$. In

particular, note that for a given precision the amount of work required to calculate $E^{(v)}(v(P_c))$ by Theorem 5 or 6 should be about the same as the amount required to calculate $f(z)$ by Corollary 4.1. Thus, in calculating $E^{(v)}(v(P_c))$, it is advantageous to express $E^{(v)}(v(P_c))$ in terms of the density function of r as in Theorems 5 and 6 rather than to express $E^{(v)}(v(P_c))$ as $\int_0^\infty z^v f(z)\,dz$.

6. EXTENSIONS

In this section we shall indicate how the preceding results can be used to obtain the distribution and/or moments of $v(P_c)$ when P_c need not satisfy all the assumptions of Section 3. However, no formal statements or proofs will be presented.

In the following, we shall refer to strengthened versions of Assumptions (2), (3), and (4) which are stated below for a stochastic geometric program P_c that satisfies Assumption (1).

We say P_c satisfies Assumption (2') iff P_c satisfies Assumption (2) and $\hat{s}^{(0)} \notin \text{span}\{\hat{s}^{(j)} | j \in <\bar{d}>\}$ where $\hat{s}^{(j)} \triangleq \beta' \hat{b}^{(j)}$ for every $j \in <\bar{d}>$. Here $\{\hat{b}^{(j)} | j \in <\bar{d}>\}$ is any nullity set for P_c and $\hat{b}^{(0)}$ is any normality vector for P_c. Also, the matrix β is defined as in Assumption (1).

P_c is said to satisfy Assumption (3') iff P_c is superconsistent and soluble for every $\bar{c} \in R^n_{++}$.

Finally, P_c is said to satisfy Assumption (4') iff D_c has a unique optimal point δ_c and $\delta_c > \underline{0}$ for every $\bar{c} \in R^n_{++}$.

Now consider a family of random cost vectors $\{c(\epsilon) | \epsilon \in (0,\infty)\}$, where $c(\epsilon) \triangleq (c_1(\epsilon), \ldots, c_n(\epsilon))'$ for $\epsilon \in (0,\infty)$, that satisfies the following:

(i) $(\log c_1(\epsilon), \ldots, \log c_n(\epsilon))'$ is a nondegenerate normal random vector;

(ii) $E(\log c_i(\epsilon)) = E(\log c_i)$ for every $i \in <n>$;

(iii) $\lim_{\epsilon \downarrow 0} \text{Cov}(\log c_i(\epsilon), \log c_j(\epsilon)) = \text{Cov}(\log c_i, \log c_j)$ for every $(i,j) \in <n> \times <n>$, where Cov denotes covariance.

Such a family of cost vectors can easily be constructed if P_c satisfies Assumption (1). When P_c also satisfies Assumptions (3') and (4') one can show that $P_{c(\epsilon)}$ will satisfy Assumptions (1), (2'), (3'), and (4'), where $P_{c(\epsilon)}$ is obtained from P_c by replacing c with the cost vector $c(\epsilon)$. Thus, the results of the preceding sections can be used to calculate the moments, distribution function, and density function of $P_{c(\epsilon)}$ for $\epsilon \in (0,\infty)$. Additionally, one can establish that the moments, distribution function, and density function of $P_{c(\epsilon)}$ converge to the corresponding moments, distribution function, and density function of P_c as ϵ tends to zero.

Next, consider the family of stochastic geometric programs $\{P_c^{(\gamma)} | \gamma \in (0,\infty)\}$ where, for $\gamma \in (0,\infty)$, $P_c^{(\gamma)}$ denotes the following stochastic program:

$$\inf_{t>\underline{0}} \sum_{i=1}^{m} c_i \prod_{j=1}^{n} t_j^{a_{ij}} \prod_{k=1}^{p} z_k^{\gamma}$$

subject to $\sum_{i \in J_\kappa} c_i \prod_{j=1}^{m} t_j^{a_{ij}} + z_\kappa \leqslant 1$ for every $\kappa \in <p>$, $t > \underline{0}$, and $z > \underline{0}$ where $t = (t_1, \ldots, t_m)'$ and $z = (z_1, \ldots, z_p)'$. One can show $P_c^{(\gamma)}$ satisfies each assumption that P_c satisfies. In addition, if P_c satisfies (3') then $P_c^{(\gamma)}$ satisfies (3') and (4') (even when P_c does not satisfy (4)). Thus, one can apply the results of the preceding sections to calculate the density function, distribution function, and moments of $P_c^{(\gamma)}$ for $\gamma \in (0, \infty)$ when P_c satisfies (1), (2'), and (3'). Furthermore, one can establish that the moments, distribution function, and density function of $P_c^{(\gamma)}$ converge to the corresponding moments, distribution function, and density function of P_c as γ approaches zero.

Finally, for $\gamma \in (0, \infty)$ and ϵ $(0, \infty)$, let $P_{c(\epsilon)}^{(\gamma)}$ denote the stochastic program obtained from $P_c^{(\gamma)}$ by replacing cost vector c by $c(\epsilon)$ in $P_c^{(\gamma)}$. The family of cost vectors $\{c(\epsilon) | \epsilon \in (0, \infty)\}$ is assumed to satisfy the properties (i), (ii), and (iii) previously listed. One can show, for $(\gamma, \epsilon) \in (0, \infty) \times (0, \infty)$, program $P_c(\epsilon)^{(\gamma)}$ satisfies (1), (2'), (3'), and (4') if P_c satisfies (1) and (3'). Thus, in this instance, one can apply the results of the preceding sections to $P_{c(\epsilon)}^{(\gamma)}$. This suggests that the family of programs $\{P_{c(\epsilon)}^{(\gamma)} | \gamma \in (0, \infty)$ and $\epsilon \in (0, \infty)\}$ may be useful in obtaining the moments, distribution function, and density function of P_c when P_c need only satisfy Assumptions (1) and (3').

APPENDIX A.

Theorem 1 in Section 3 is an immediate consequence of the following lemma.

LEMMA A: Define $L(z,s) \triangleq \sum_{i=1}^{u} s_i \log z_i$ for every positive-valued random vector $z \triangleq (z_1, \ldots, z_u)'$ and $s \in R^u$. Also, define the inner product $<\cdot,\cdot>_\Lambda$ on R^u by $<x,y>_\Lambda \triangleq x'\Lambda y$ for $(x,y) \in R^u \times R^u$. (Note $<\cdot,\cdot>_\Lambda$ is an inner product since Λ is a dispersion matrix of a nondegenerate normal random vector and hence is positive definite.) Then

(i) $(L(e,s^{(1)}), \ldots, L(e,s^{(d)}))'$ is a normal random vector with independent components where $e \triangleq (e_1, \ldots, e_u)'$, $s^{(1)} = \hat{s}^{(1)}$, and

$$s^{(j)} = \hat{s}^{(j)} - \sum_{i=1}^{j-1} (<s^{(i)}, s^{(i)}>_\Lambda)^{-1} (<\hat{s}^{(j)}, s^{(i)}>_\Lambda) s^{(i)}$$

for $1 < j \leqslant d$;

(ii) $(L(e,s^{(0)}), L(e,s^{(1)}), \ldots, L(e,s^{(d)}))'$ is a normal random vector with independent components if $\hat{s}^{(0)} \notin$ span $\{\hat{s}^{(j)} | j \in <d>\}$ where $s^{(0)} = \hat{s}^{(0)} - \sum_{i=1}^{d} (<s^{(i)}, s^{(i)}>_\Lambda)^{-1} (<\hat{s}^{(0)}, s^{(i)}>_\Lambda) s^{(i)}$ when $\hat{s}^{(0)} \notin$ span $\{\hat{s}^{(j)} | j \in <d>\}$.

(iii) For every $j \in <\bar{d}>$, $s^{(j)} = \beta' b^{(j)}$ and $\prod_{i=1}^{n} c_i^{b_i^{(j)}} = \left[\prod_{i=1}^{n} \alpha_i^{b_i^{(j)}}\right] \{\exp(L(e,s^{(j)}))\}$ where $b^{(1)} = \hat{b}^{(1)}$ and, for $1 < j \leqslant d$, $b^{(j)} = \hat{b}^{(j)} - \sum_{i=1}^{j-1} (<\beta' b^{(i)}, \beta' b^{(i)}>_\Lambda)^{-1} (<\beta' \hat{b}^{(j)}, \beta' b^{(i)}>_\Lambda) b^{(i)}$. Also $b^{(0)} = \hat{b}^{(0)}$ if $\hat{s}^{(0)} \in$ span $\{\hat{s}^{(j)} | j \in <d>\}$ and $b^{(0)} = \hat{b}^{(0)} - \sum_{i=1}^{d} (<\beta' b^{(i)}, \beta' b^{(i)}>_\Lambda)^{-1} (<\beta' b^{(0)}, \beta' b^{(i)}>_\Lambda) b^{(i)}$ if $\hat{s}^{(0)} \notin$ span $\{\hat{s}^{(j)} | j \in <d>\}$. Furthermore, $\{b^{(j)} | j \in <d>\}$ is a nullity set and $b^{(0)}$ is a normality vector of P_c;

(iv) For $j \in <\bar{d}>$, the density function ϕ_j of $L(e,s^{(j)})$ is given by $\phi_j(l) \triangleq \frac{1}{\omega_j\sqrt{2\pi}} \exp\left[-\frac{(l-\rho_j)^2}{2\omega_j^2}\right]$ for every $l \in R^1$ where ρ_j is the expected value of $L(e,s^{(j)})$ and ω_j^2 is the variance of $L(e,s^{(j)})$. Furthermore, $\rho_j = \sum_{i=1}^{u} \mu_i s_i^{(j)}$ and $\omega_j^2 = <s^{(j)},s^{(j)}>_\Lambda$.

PROOF: Since Λ is real symmetric, Λ has an orthonormal set of u eigenvectors $\{p_1, \ldots, p_u\}$. Let P be the $u \times u$ matrix whose jth column is p_j. Then P is orthogonal (i.e., $P^{-1} = P'$) and $\tilde{\Lambda} \triangleq P^{-1}\Lambda P$ is diagonal.

For every $i \in <u>$ let $\gamma_i \triangleq \log e_i$ and $\gamma \triangleq (\gamma_1, \ldots, \gamma_u)'$. Let $\tilde{\gamma} \triangleq P'\gamma$. Then $\tilde{\gamma}$ is a u-variate normal vector with dispersion matrix $P'\Lambda P = \tilde{\Lambda}$ ([8, Th. 2.1.1]). Since $\tilde{\Lambda}$ is diagonal, the components of $\tilde{\gamma}$ are independent.

Let $(s,w) \in R^u \times R^u$ such that $w = P's$. We shall show $L(e,s) = L(\tilde{e},w)$ where $\tilde{e} \triangleq (\tilde{e}_1, \ldots, \tilde{e}_u)'$ and $\tilde{e} \triangleq e^{\gamma}$ for $i \in <u>$. Note $\gamma = P\tilde{\gamma}$. Thus, for $i \in <u>$, $\gamma_i = \sum_{k=1}^{u} p_{ik}\tilde{\gamma}_k$. Hence,

$$(1) \qquad L(e,s) = \sum_{i=1}^{u} s_i \left[\sum_{k=1}^{u} p_{ik}\tilde{\gamma}_k\right] = \sum_{k=1}^{u} \left[\sum_{i=1}^{u} p_{ik}s_i\right]\tilde{\gamma}_k$$

$$= \sum_{k=1}^{u} w_k\tilde{\gamma}_k = L(\tilde{e},w).$$

For every $j \in <d>$ define $\hat{w}^{(j)} \triangleq P's^{(j)}$. By assumption $\{\hat{s}^{(j)} | j \in <d>\}$ is linearly independent. Thus $\{\hat{w}^{(j)} | j \in <d>\}$ is linearly independent since P' is nonsingular. Thus, one can apply the Gram-Schmidt orthogonalization process to $\{\hat{w}^{(j)} | j \in <d>\}$ to obtain the orthogonal set $\{w^{(j)} | j \in <d>\}$ with respect to $<\cdot,\cdot>_\Lambda$ where $w^{(1)} \triangleq \hat{w}^{(1)}$ and, for $1 < j \leqslant d$,

$$(2) \qquad w^{(j)} \triangleq \hat{w}^{(j)} - \sum_{l=1}^{j-1} (<w^{(l)},w^{(l)}>_\Lambda)^{-1}(<\hat{w}^{(j)},w^{(l)}>_\Lambda)w^{(l)}.$$

(Note $<\cdot,\cdot>_\Lambda$ is the inner product on R^u defined by $<x,y>_\Lambda \triangleq x'\tilde{\Lambda}y$ for $(x,y) \in R^u \times R^u$. $<\cdot,\cdot>_\Lambda$ is an inner product on R^u since $\tilde{\Lambda}$ is the dispersion matrix of the nondegenerate normal random vector $\tilde{\gamma}$ and hence is positive definite.) Also define

$$(3) \qquad w^{(0)} \triangleq \hat{w}^{(0)} - \sum_{l=1}^{d} (<w^{(l)},w^{(l)}>_\Lambda)^{-1}(<\hat{w}^{(0)},w^{(l)}>_\Lambda)w^{(l)}$$

if $\hat{s}^{(0)} \notin$ span $\{\hat{s}^{(j)} | j \in <d>\}$; otherwise define $w^{(0)} \triangleq \hat{w}^{(0)}$. Observe if $\hat{s}^{(0)} \notin$ span $\{\hat{s}^{(j)} | j \in <d>\}$ then $\hat{w}^{(0)} \notin$ span $\{\hat{w}^{(j)} | j \in <d>\}$. Thus, $\{w^{(j)} | j \in <\bar{d}>\}$ is an orthogonal set in R^u with respect to $<\cdot,\cdot>_\Lambda$ when $\hat{s}^{(0)} \notin$ span $\{\hat{s}^{(j)} | j \in <d>\}$.

Define $s^{(j)} \triangleq Pw^{(j)}$ for every $j \in <\bar{d}>$. Then, for $j \in <\bar{d}>$, $w^{(j)} = P's^{(j)}$. Thus, by (1),

$$(4) \qquad L(e,s^{(j)}) = L(\tilde{e},w^{(j)}) \text{ for every } j \in <\bar{d}>.$$

We shall next show $(L(\tilde{e},w^{(1)}), \ldots, L(\tilde{e},w^{(d)}))'$ is a normal random vector with independent components. Also, whenever $\hat{s}^{(0)} \notin$ span $\{\hat{s}^{(j)} | j \in <d>\}$, we shall show that $(L(\tilde{e},w^{(0)}), L(\tilde{e},w^{(1)}), \ldots, L(\tilde{e},w^{(d)}))'$ is a normal random vector with independent components.

For every $i \in <u>$ let θ_i^2 and τ_i be the variance and mean, respectively, of $\log \tilde{e}_i$. For $i \in <u>$ define $\psi_i \triangleq \theta_i^{-1}(\log \tilde{e}_i - \tau_i)$. Note, for every $j \in <\bar{d}>$, $L(\tilde{e},w^{(j)}) = \sum_{i=1}^{u} w_i^{(j)} \log \tilde{e}_i = \sum_{i=1}^{u} \theta_i w_i^{(j)} \psi_i + \sum_{i=1}^{u} \tau_i w_i^{(j)}$. Let $r,t \in <\bar{d}>$ such that $r \neq t$. Recall $\tilde{\gamma} \triangleq (\log \tilde{e}_1, \ldots, \log \tilde{e}_u)'$ is a normal vector with independent components. Thus, $\{\psi_i | i \in <u>\}$ is a set of independent unit normal random variables. Thus, $L(\tilde{e},w^{(r)})$ and $L(\tilde{e},w^{(t)})$ are normal random variables. Moreover, $L(\tilde{e},w^{(r)})$ and $L(\tilde{e},w^{(t)})$ are independent provided $\sum_{i=1}^{u} \theta_i^2 w_i^{(r)} w_i^{(t)} = 0$ [8, Th. 4.1.1, p. 70].

Since, for every $i \in <u>$, θ_i^2 is the variance of $\tilde{\gamma}_i$ and $\tilde{\Lambda}$ is the dispersion matrix of $\tilde{\gamma}$ one has $\sum_{i=1}^{u} \theta_i^2 w_i^{(r)} w_i^{(t)} = <w^{(r)},w^{(t)}>_{\tilde{\Lambda}}$. By construction of $\{w^{(j)} | j \in <\bar{d}>\}$ one has $<w^{(r)},w^{(t)}>_{\tilde{\Lambda}} = 0$ for $r,t \in <d>$ with $r \neq t$. Also, $<w^{(r)},w^{(t)}>_{\tilde{\Lambda}} = 0$ for $r,t \in <\bar{d}>$ with $r \neq t$ provided $\hat{s}^{(0)} \notin$ span $\{\hat{s}^{(j)} | j \in <d>\}$. Hence, by (4), $(L(e,s^{(1)}), \ldots, L(e,s^{(d)}))'$ is a normal random vector with independent components. Also by (4), $(L(e,s^{(0)}), L(e,s^{(1)}), \ldots, L(e,s^{(d)}))'$ is a normal random vector with independent components if $\hat{s}^{(0)} \notin$ span $\{\hat{s}^{(j)} | j \in <d>\}$.

Next, let $(x^{(i)},y^{(i)}) \in R^u \times R^u$ for $i \in \{1,2\}$ such that $y^{(i)} = P'x^{(i)}$. Then,

$$<y^{(1)},y^{(2)}>_{\tilde{\Lambda}} = <P'x^{(1)},P'x^{(2)}>_{\tilde{\Lambda}} = (x^{(1)})'P\tilde{\Lambda}P'x^{(2)} = <x^{(1)},x^{(2)}>_{\Lambda}. \tag{5}$$

Observe $s^{(1)} = Pw^{(1)} = P\hat{w}^{(1)} = \hat{s}^{(1)}$. Also, by (2) and (5), for $1 < j \leqslant d$ one has

$$\begin{aligned} s^{(j)} &= Pw^{(j)} = P\hat{w}^{(j)} - \sum_{l=1}^{j-1} (<w^{(l)},w^{(l)}>_{\tilde{\Lambda}})^{-1} (<\hat{w}^{(j)},w^{(l)}>_{\tilde{\Lambda}}) Pw^{(l)} \\ &= \hat{s}^{(j)} - \sum_{l=1}^{j-1} (<s^{(l)},s^{(l)}>_{\Lambda})^{-1} (<\hat{s}^{(j)},s^{(l)}>_{\Lambda}) s^{(l)}. \end{aligned} \tag{6}$$

Moreover, by (3) and (5) one has

$$s^{(0)} = Pw^{(0)} = \hat{s}^{(0)} - \sum_{l=1}^{d} (<s^{(l)},s^{(l)}>_{\Lambda})^{-1} (<\hat{s}^{(0)},s^{(l)}>_{\Lambda}) s^{(l)} \tag{7}$$

if $\hat{s}^{(0)} \notin$ span $\{\hat{s}^{(j)} | j \in <d>\}$. This completes the demonstration of (i) and (ii).

Next, we shall obtain a nullity set $\{b^{(j)} | j \in <d>\}$ and normality vector $b^{(0)}$ for P_c such that $s^{(j)} = \beta' b^{(j)}$ and $\prod_{i=1}^{n} c_i^{b_i^{(j)}} = \left\{\prod_{i=1}^{n} \alpha_i^{b_i^{(j)}}\right\} \{\exp(L(e,s^{(j)}))\}$ for every $j \in <\bar{d}>$. Let $S \triangleq$ span $\{\hat{s}^{(j)} | j \in <d>\}$ and $B \triangleq$ span $\{\hat{b}^{(j)} | j \in <d>\}$. By assumption $\{\hat{s}^{(j)} | j \in <d>\}$ is linearly independent and hence a basis for S. Thus, there exists a unique linear transformation $T: S \to B$ such that $T(\hat{s}^{(j)}) = \hat{b}^{(j)}$ for every $j \in <d>$. Since $\{\hat{b}^{(j)} | j \in <d>\}$ is a basis for B, T is an isomorphism from S onto B. Also, since $T^{-1}(\hat{b}^{(j)}) = \hat{s}^{(j)} = \beta'\hat{b}^{(j)}$ for every $j \in <d>$ one has $T^{-1}(b) = \beta' b$ for every $b \in B$.

Recall $s^{(1)} = \hat{s}^{(1)} \in S$. Let $1 < j \leqslant d$ and assume $s^{(l)} \in S$ for $1 \leqslant l < j$. Then by (6) one has $s^{(j)} \in S$. Hence, $\{s^{(j)} | j \in <d>\} \subset S$. Also $\{s^{(j)} | j \in <d>\}$ is linearly independent since $\{w^{(j)} | j \in <d>\}$ is orthogonal with respect to $<\cdot,\cdot>_{\Lambda}$ and P is a nonsingular matrix for

which $s^{(j)} = Pw^{(j)}$ for every $j \in <d>$. Thus, $\{s^{(j)} | j \in <d>\}$ is a basis for S. Since T is an isomorphism from S onto B, $\{T(s^{(j)}) | j \in <d>\}$ is a basis for B. Thus, $\{b^{(j)} | j \in <d>\}$ is a nullity set for P_c where $b^{(j)} \underline{\Delta} T(s^{(j)})$ for $j \in <d>$.

Let $\tilde{S} \underline{\Delta}$ span $\{\hat{s}^{(j)} | j \in <\bar{d}>\}$ and $\tilde{B} \underline{\Delta}$ span $\{\hat{b}^{(j)} | j \in <\bar{d}>\}$. Suppose $\hat{s}^{(0)} \notin S$. Then there exists a unique linear transformation $\tilde{T}: \tilde{S} \to \tilde{B}$ such that $\tilde{T}(\hat{s}^{(j)}) = \hat{b}^{(j)}$ for every $j \in <\bar{d}>$. Also, by (7), $s^{(0)} \in \tilde{S}$. Thus, we can define $b^{(0)} \underline{\Delta} \tilde{T}(s^{(0)})$. Note by the definition of $\tilde{T}$ one has $\tilde{T}(s) = T(s)$ for every $s \in S$. Thus, by (7),

$$(8) \quad b^{(0)} = \tilde{T}(\hat{s}^{(0)}) - \sum_{l=1}^{d} (<s^{(l)}, s^{(l)}>_1)^{-1} (<\hat{s}^{(0)}, s^{(l)}>_1) \tilde{T}(s^{(l)})$$

$$= \hat{b}^{(0)} - \sum_{l=1}^{d} (<s^{(l)}, s^{(l)}>_1)^{-1} (<\hat{s}^{(0)}, s^{(l)}>_1) b^{(l)}$$

since $\tilde{T}(s^{(l)}) = T(s^{(l)}) = b^{(l)}$ for $l \in <d>$. For $j \in <m>$ let A_j denote column j of exponent matrix A. Recall $q \in R^n$ such that $q_i = 1$ if $i \in <n_0>$ and $q_i = 0$ if $i > n_0$ where n_0 is the number of elements in J_0. Then by (8), for every $j \in <m>$, one has $<b^{(0)}, A_j> = 0$ since $\hat{b}^{(0)}$ is a normality vector and $\{b^{(l)} | l \in <d>\}$ is a nullity set of P_c, where $<\cdot, \cdot>$ denotes the usual inner product on R^n. Also, $<b^{(0)}, q> = <\hat{b}^{(0)}, q> = 1$ since $<b^{(l)}, q> = 0$ for every $l \in <d>$. Thus, $b^{(0)}$ is a normality vector for P_c. If $\hat{s}^{(0)} \in S$ we define $b^{(0)} \underline{\Delta} \hat{b}^{(0)}$. Thus, whether $\hat{s}^{(0)} \notin S$ or $\hat{s}^{(0)} \in S$ one has that $b^{(0)}$ is a normality vector for P_c.

To show $\beta' b^{(j)} = s^{(j)}$ for every $j \in <\bar{d}>$ first observe for $j \in <d>$ one has $\beta' b^{(j)} = T^{-1}(b^{(j)}) = T^{-1}(T(s^{(j)}) = s^{(j)}$. Next suppose $\hat{s}^{(0)} \notin S$. Then $\tilde{T}$ is an isomorphism from $\tilde{S}$ onto $\tilde{B}$ since $\{\hat{b}^{(j)} | j \in <\bar{d}>\}$ is linearly independent and $\tilde{T}(\hat{s}^{(j)}) = \hat{b}^{(j)}$ for every $j \in <\bar{d}>$. Also, $\tilde{T}^{-1}(\hat{b}^{(j)}) = \hat{s}^{(j)} = \beta' \hat{b}^{(j)}$ for every $j \in <\bar{d}>$. Hence, $\tilde{T}^{-1}(b) = \beta' b$ for all $b \in B$. Note $b^{(0)} \underline{\Delta} \tilde{T}(s^{(0)}) \in B$. Thus, $\beta' b^{(0)} = \tilde{T}^{-1}(b^{(0)}) = \tilde{T}^{-1}(\tilde{T}(s^{(0)})) = s^{(0)}$. Finally, suppose $\hat{s}^{(0)} \in S$. Then $B' b(0) = \beta' \hat{b}^{(0)} = \hat{s}^{(0)} = P\hat{w}^{(0)} = Pw^{(0)} = s^{(0)}$. Thus, from the above, $\beta' b^{(j)} = s^{(j)}$ for every $j \in <\bar{d}>$.

Next let $j \in <\bar{d}>$ and observe $\prod_{i=1}^{n} c_i^{b_i^{(j)}} = \prod_{i=1}^{n} \left[\alpha_i \prod_{\nu=1}^{u} e_\nu^{\beta_{i\nu}}\right]^{b_i^{(j)}} = \left[\prod_{i=1}^{n} \alpha_i^{b_i^{(j)}}\right] \left\{\prod_{\nu=1}^{u} \left[\prod_{i=1}^{n} e_\nu^{\beta_{i\nu} b_i^{(j)}}\right]\right\} = \left[\prod_{i=1}^{n} \alpha_i^{b_i^{(j)}}\right] \left\{\prod_{\nu=1}^{u} e_\nu^{\sum_{i=1}^{n} \beta_{i\nu} b_i^{(j)}}\right\}$. Thus, since $\beta' b^{(j)} = s^{(j)}$, one has $\prod_{i=1}^{n} c_i^{b_i^{(j)}} = \left[\prod_{i=1}^{n} \alpha_i^{b_i^{(j)}}\right] \left\{\prod_{\nu=1}^{u} e_\nu^{s_\nu^{(j)}}\right\} = \left[\prod_{i=1}^{n} \alpha_i^{b_i^{(j)}}\right] \left\{\exp\left[\sum_{\nu=1}^{u} s_\nu^{(j)} \log e_\nu\right]\right\} = \left[\prod_{i=1}^{n} \alpha_i^{b_i^{(j)}}\right] \{\exp(L(e, s^{(j)}))\}$.

Note $b^{(1)} = T(s^{(1)}) = T(Pw^{(1)}) = T(P\hat{w}^{(1)}) = T(\hat{s}^{(1)}) = \hat{b}^{(1)}$. Also, by (6), for $1 < j \leqslant d$ one has $b^{(j)} = T(s^{(j)}) = T(\hat{s}^{(j)}) - \sum_{l=1}^{j-1} (<s^{(l)}, s^{(l)}>_1)^{-1} (<\hat{s}^{(j)}, s^{(l)}>_1) T(s^{(l)}) = b^{(j)} - \sum_{l=1}^{j-1} (<\beta' b^{(l)}, \beta' b^{(l)}>_1)^{-1} (<\beta' b^{(j)}, \beta' b^{(l)}>_1) b^{(l)}$. Recall if $\hat{s}^{(0)} \in$ span $\{\hat{s}^{(j)} | j \in <d>\}$ then $b^{(0)} \underline{\Delta} \hat{b}^{(0)}$. If $\hat{s}^{(0)} \notin$ span $\{\hat{s}^{(j)} | j \in <d>\}$ then by (8) one has

$$b^{(0)} = \hat{b}^{(0)} - \sum_{l=1}^{d} (<\beta' b^{(l)}, \beta' b^{(l)}>_1)^{-1} (<\beta' \hat{b}^{(0)}, \beta' b^{(l)}>_1) b^{(l)}.$$

This completes the demonstration of (iii).

Finally, let $j \in <\bar{d}>$ and recall $L(e,s^{(j)}) = \sum_{i=1}^{u} s_i^{(j)} \log e_i$. Thus, $\rho_j = \sum_{i=1}^{u} s_i^{(j)}\mu_i$ and $\omega_j^2 = <s^{(j)}, s^{(j)}>_\Lambda$ [8, Th. 2.1.1, p. 29] since $(\mu_1, \ldots, \mu_u)'$ is the mean vector and Λ is the dispersion matrix of e. Also, since $L(e,s^{(j)})$ is normal, one has $\phi_j(l) = \frac{1}{\omega_j\sqrt{2\pi}} \exp\left[-\frac{(l-\rho_j)^2}{2\omega_j^2}\right]$ for every $l \in R^1$.

APPENDIX B.

PROPOSITION B: (a) If H is bounded then $E^{(\nu)}(v(P_c))$ exists for every $\nu \in N$.

(b) H is bounded iff the set $F \triangleq \{\delta \in R^n | \delta_i \geqslant 0 \ \forall\ i \in <n>,\ A'\delta = \underline{0}, \text{ and } q'\delta = 1\}$ is bounded.

PROOF: (a) Assume H is bounded. Then for every $j \in <d>$ there exist real numbers l_j and u_j such that $l_j \leqslant r \leqslant u_j$ for every $r \in H$. Let $\nu \in N$ and define $z_c \triangleq v(P_c)$. Finally, assume c_j takes on the value $\bar{c}_j$ and recall $r_c = (r_1(\bar{c}), \ldots, r_d(\bar{c}))'$ is the element of H for which $\hat{\delta} \triangleq b^{(0)} + \sum_{j=1}^{d} (r_j(\bar{c}))b^{(j)}$ is the unique optimal point of D_c.

Since $\hat{\delta}$ is the optimal point of D_c, by [4, Ch. 3, Sec. 3] and Assumption 3 one has

$$z_c^\nu = K_0(\bar{c})^\nu \prod_{j=1}^{d} K_j(\bar{c})^{\nu r_j(\bar{c})} \prod_{i=1}^{n} \delta_i(r_c)^{\nu\delta_i(r_c)} \prod_{\kappa=1}^{p} \lambda_\kappa(r_c)^{\nu\lambda_\kappa(r_c)}$$

where $\prod_{\kappa=1}^{p} \lambda_\kappa(r)^{\lambda_\kappa(r)} \triangleq 1$ for $r \in \bar{H}$, the closure of H, if $p = 0$. Define τ: $\bar{H} \to R^1$ by $\tau(r) \triangleq \prod_{i=1}^{n} \delta_i(r)^{\delta_i(r)} \prod_{\kappa=1}^{p} \lambda_\kappa(r)^{\lambda_\kappa(r)}$ for $r \in \bar{H}$. In evaluating $\tau(r)$ use the convention $x^1 = x^{-1} = 1$ for $x = 0$. Then τ is continuous on $\bar{H}$. Thus, since $\bar{H}$ is compact, there exists $U \in (0,\infty)$ such that $0 < \tau(r) \leqslant U$ for every $r \in H$. Hence,

$$0 < z_c^\nu \leqslant U^\nu K_0(\bar{c})^\nu \prod_{j=1}^{d} K_j(\bar{c})^{\nu r_j(\bar{c})}. \tag{1}$$

I. Assume $\xi^{(0)} \in \text{span}\{\xi^{(j)} | j \in <d>\}$. Then by Proposition 3 there exists $y_j \in R^1$ for $j \in <d>$ and $W \in (0,\infty)$ such that $K_0(c) = W \prod_{j=1}^{d} K_j(c)^{y_j}$. Thus, by (1),

$$0 < z_c^\nu \leqslant (UW)^\nu \prod_{j=1}^{d} K_j(\bar{c})^{\nu(y_j + r_j(\bar{c}))}. \tag{2}$$

Let $j \in <d>$. If $0 < K_j(\bar{c}) < 1$ then $K_j(\bar{c})^{\nu y_j} K_j(\bar{c})^{\nu r_j(\bar{c})} \leqslant K_j(\bar{c})^{\nu y_j} K_j(\bar{c})^{\nu l_j}$. Also, if $1 \leqslant K_j(\bar{c})$ then $K_j(\bar{c})^{\nu y_j} K_j(\bar{c})^{\nu r_j(\bar{c})} \leqslant K_j(\bar{c})^{\nu y_j} K_j(\bar{c})^{\nu u_j}$. Thus, $0 < K_j(\bar{c})^{\nu(y_j + r_j(\bar{c}))} \leqslant Z_j(\bar{c}) \triangleq \max(K_j(\bar{c})^{\nu(y_j + l_j)}, K_j(\bar{c})^{\nu(y_j + u_j)})$. Hence, by (2),

$$0 < z_c^\nu \leqslant (UW)^\nu \prod_{j=1}^{d} Z_j(\bar{c}). \tag{3}$$

Moreover, by the choice of $\{b^{(j)} | j \in <d>\}$, the variates $K_j(c)$ for $j \in <d>$ are independent and hence $\{Z_j(c) | j \in <d>\}$ is a set of independent variates.

By definition of $Z_j(\bar{c})$ one has

$$0 < Z_j(\bar{c}) \leqslant K_j(\bar{c})^{r(v_j+l_j)} + K_j(\bar{c})^{r(v_j+u_j)}. \tag{4}$$

Since $K_j(c)$ is lognormal so are $K_j(c)^{r(v_j+l_j)}$ and $K_j(c)^{r(v_j+u_j)}$. Thus, the expected values of these two variates exist and hence so does $E(K_j(c)^{r(v_j+l_j)} + K_j(c)^{r(v_j+u_j)})$. Thus, by (4), the expected value of $Z_j(c)$ exists. Since the variates $Z_j(c)$ for $j \in <d>$ are independent the expected value of $(UW)^r \prod_{j=1}^{d} Z_j(c)$ must also exist [5, p. 82, Th. 3.6.2]. Hence, by (3), the expected value of z_c^r exists.

II. Assume $\delta^{(0)} \notin \text{span}\{\delta^{(j)} | j \in <d>\}$. Then by the choice of $\{b^{(j)} | j \in <\bar{d}>\}$ the variates $K_j(c)$ for $j \in <\bar{d}>$ are independent. For $j \in <d>$ define $Z_j(\bar{c}) \triangleq \max(K_j(\bar{c})^{rl_j}, K_j(\bar{c})^{ru_j})$ and let $Z_0(\bar{c}) \triangleq U^r K_0(\bar{c})^r$. Then the variates $Z_j(c)$ for $j \in <\bar{d}>$ must be independent. Furthermore, $0 < K_j(\bar{c})^{rv_j(c)} \leqslant Z_j(\bar{c})$ for $j \in <d>$. Hence, by (1),

$$0 < z_c^r \leqslant \prod_{j=0}^{d} Z_j(\bar{c}). \tag{5}$$

Note that $E(Z_0(c))$ exists since $Z_0(c)$ is lognormal. Also, for $j \in <d>$, $E(Z_j(c))$ exists since $K_j(c)$ is lognormal and $0 < Z_j(\bar{c}) < K_j(\bar{c})^{rl_j} + K_j(\bar{c})^{ru_j}$. Since the variates $Z_j(c)$ for $j \in <\bar{d}>$ are independent it follows from (5) that $E(z_c^r)$ exists.

By I and II we conclude $E^{(c)}(v(P_c))$ exists for all $r \in N$.

(b) Observe $\delta \in F$ iff there exists $r = (r_1, \ldots, r_d)' \in \bar{H}$ such that $\delta = b^{(0)} + \sum_{j=1}^{d} r_j b^{(j)}$. Also, $\bar{H}$ is bounded iff H is bounded. Thus, F is bounded iff H is bounded.

REFERENCES

[1] Abrams, R.A., "Consistency, Superconsistency, and Dual Degeneracy in Posynomial Geometric Programming," Operations Research, *24*, 325-335 (1976).

[2] Aitchison, J. and J.A.C. Brown, *The Lognormal Distribution*, (Cambridge University Press, London, England, 1957).

[3] Avriel, M. and D.J. Wilde, "Stochastic Geometric Programming," Proceedings Princeton Symposium on Mathematical Programming, Princeton, New Jersey (1970).

[4] Duffin, R.J., E.L. Peterson and C. Zener, *Geometric Programming*, (John Wiley & Sons, New York, New York, 1967).

[5] Fisz, M., *Probability Theory and Mathematical Statistics*, (John Wiley & Sons, New York, New York, 1963).

[6] Hammersley, J.M. and D.C. Handscomb, *Monte Carlo Methods*, (John Wiley & Sons, New York, New York, 1964).

[7] Isaacson, E. and H.B. Keller, *Analysis of Numerical Methods*, (John Wiley & Sons, New York, New York, 1966).

[8] Lukacs, E. and R.G. Laha, "*Applications of Characteristic Functions*, (Charles Griffin & Company, Ltd., London, England, 1964).

[9] McNichols, G.R., "On the Treatment of Uncertainty in Parametric Costing," Ph.D. Thesis, The George Washington University, Washington, D.C. (1976).

[10] Munroe, M.E., *Introduction to Measure and Integration*, (Addison-Wesley Publishing Company, Reading, Massachusetts, 1953).

[11] Rotar, V.I., "On the Speed of Convergence in the Multidimensional Central Limit Theorem," Theory of Probability and Its Applications, *15*, 354-356 (1970).

[12] Spivak, M., *Calculus on Manifolds*, (W.A. Benjamin, New York, New York, 1965).

[13] Stark, R.M., "On Zero-Degree Stochastic Geometric Programs," Journal of Optimization Theory and Applications, *23*, 167-187 (1977).

[14] Tsuda, T., "Numerical Integration of Functions of Very Many Variables," Numerische Mathematik, *20*, 377-391 (1973).

[15] Turchin, V.F., "On the Computation of Multidimensional Integrals by the Monte-Carlo Method," Theory of Probability and Its Applications, *16*, 720-724 (1971).

A CLASS OF CONTINUOUS NONLINEAR PROGRAMMING PROBLEMS WITH TIME-DELAYED CONSTRAINTS

Thomas W. Reiland

North Carolina State University
Raleigh, North Carolina

Morgan A. Hanson

Florida State University
Tallahassee, Florida

ABSTRACT

A general class of continuous time nonlinear problems is considered. Necessary and sufficient conditions for the existence of solutions are established and optimal solutions are characterized in terms of a duality theorem. The theory is illustrated by means of an example.

1. INTRODUCTION

Recently Farr and Hanson [1] proved existence theorems, duality theorems, and continuous time analogues of the Kuhn-Tucker Theorem for a class of continuous time programming problems in which nonlinearity appears both in the objective function and in the constraints. More recently this class was extended in Farr and Hanson [2] to include problems with prescribed time lags in the constraints. In this paper we generalize these results by considering a more general form of the constraints and by assuming a less stringent constraint qualification. This constraint qualification is analogous to that of Kuhn and Tucker [5] and provides further unification between the areas of finite-dimensional and continuous time programming. An example is presented wherein these results are applied to a version of Koopmans' [4] water storage problem which has been modified to address the economic ramifications of an energy crisis.

2. THE PRIMAL PROBLEM

The problem under consideration (Primal Problem A) is:

Maximize

$$I(z) = \int_0^T \phi(z(t),t)\,dt$$

subject to the constraints

(1) $z(t) \geq 0, \quad 0 \leq t \leq T,$

(2) $f(z(t),t) \leq h(x(z,t),t), \quad 0 \leq t \leq T,$

and

$$z(t) = 0, \quad t < 0, \tag{3}$$

where $z \in L_n^\infty [0, T]$, i.e., z is a bounded and measurable n-dimensional function; y is a mapping from $L_n^\infty [0, T] \times [0, T]$ into E^p defined by

$$y(z,t) = \sum_{j=0}^{r} \int_0^t g_j(z(s - \alpha_j), s - \alpha_j) ds; \tag{4}$$

$f(z(t),t)$, $h(y(z,t),t) \in E^m$; $g_j(z(s - \alpha_j), s - \alpha_j) \in E^p$, $j = 0, \ldots, r$. The set $0 = \alpha_0 \leqslant \alpha_1 \leqslant \ldots \leqslant \alpha_r$ is a finite collection of nonnegative numbers; and ϕ is a scalar function, concave and continuously differentiable in its first argument throughout $[0, T]$.

It is further assumed that each component of $-f$, g_j, and h is a scalar function, concave and differentiable in its first argument throughout $[0,T]$, that each component of the composite function $h(y(\cdot,t),t): L_n^\infty [0,T] \to E^m$ is concave in z, there exists $\delta > 0$ such that

$$\text{either } \nabla_k f_i(\eta,t) = 0 \text{ or } \nabla_k f_i(\eta,t) \geqslant \delta, \tag{5}$$

and for each t and k there exists $i_k = i_k(t)$ such that

$$\nabla_k f_{i_k}(\eta,t) \geqslant \delta, \tag{6}$$

where

$$\nabla_k f_i(\eta,t) = \partial f_i(\eta,t)/\partial \eta_k,$$

$$i = 1, \ldots, m, \quad k = 1, \ldots, n,$$

for $\eta \in E^n$, $\eta \geqslant 0$, and $t \in [0,T]$;

$$g_j(z(t),t) = 0, \quad t < 0, \tag{7}$$

$$j = 0, \ldots, r;$$

$$\nabla_h h_i(v,t) = \partial h_i(v,t)/\partial v_h \geqslant 0, \tag{8}$$

for $v \in E^p$ and $t \in [0,T]$; and

$$\sup_{0 < t \leqslant T} h_i(0,T) < \infty, \quad \sup_{0 < t < T} \nabla_k h_i(0,t) < \infty, \quad i=1, \ldots, m, \tag{9a}$$

$$k = 1, \ldots, n,$$

$$\sup_{0 < t < T} g_{jq}(0,t) < \infty, \quad \sup_{0 < t < T} \nabla_k g_{jq}(0,t) < \infty, \quad j = 0, \ldots, r, \tag{9b}$$

$$q = 1, \ldots, p, \quad k = 1, \ldots, n,$$

$$\inf_{0 < t < T} f_i(0,t) > -\infty, \quad i = 1, \ldots, m, \tag{9c}$$

$$\sup_{0 < t < T} \nabla_k \phi(\eta,t) < \infty, \quad \eta \in E^n, \ \eta \geqslant 0, \ k = 1, \ldots, n. \tag{9d}$$

A function $z \in L_n^\infty [0,T]$ is termed *feasible* for Primal Problem A if it satisfies the constraints (1), (2), and (3). The primal problem is itself said to be feasible if a feasible z exists.

It should be noted that Primal Problem A is identical to that considered in [2] if $p = m$ and

$$h(y(z,t),t) = I_m y(z,t)$$

where I_m is an m-dimensional identity matrix.

3. EXISTENCE THEOREM

THEOREM 1: If Primal Problem A is feasible, then it has an optimal solution, that is, there exists a feasible $\bar{z}$ for which

$$V(\bar{z}) = \sup V(z),$$

where the supremum is taken over all feasible z.

We preface the proof of this theorem with a brief discussion of weak convergence and two lemmas.

Let X be a normed linear space and denote by X^* the collection of all bounded linear functionals on X. If we define the norm of an element $f \in X^*$ by

$$||f|| = \sup_{||x||=1} |f(x)|$$

and define addition and scalar multiplication of linear functionals in the obvious manner, then X^* is a Banach space and is commonly referred to as the dual space of X. A sequence $\{x_n\}$ in X is said to converge weakly to $x \in X$ if $f(x_n) \to f(x)$ as $n \to \infty$, for every $f \in X^*$.

LEMMA 1: Let the uniformly bounded sequence of scalar measurable functions $\{q_d(t)\}$, $d = 1, 2, \ldots$, converge weakly on $[0, T]$ to $q_0(t)$. Then except on a set of measure zero

$$q_0(t) \leqslant \lim_{d \to \infty} \sup q_d(t).$$

PROOF: See Levinson [6]

LEMMA 2: If q is a nonnegative integrable function for which there exists scalar constants $\theta_1 \geqslant 0$ and $\theta_2 > 0$ such that

$$q(t) \leqslant \theta_1 + \theta_2 \int_0^t q(s)ds, \quad 0 \leqslant t \leqslant T,$$

then $q(t) \leqslant \theta_1 e^{\theta_2 t}, \quad 0 \leqslant t \leqslant T.$

PROOF: See Levinson [6].

PROOF OF THEOREM 1: Let z be feasible for Primal Problem A and multiply the constraint (2) by the m-dimensional vector $(1, \ldots, 1)$ to obtain the inequality

$$\sum_{i=1}^{m} f_i(z(t), t) \leqslant \sum_{i=1}^{m} h_i(y(z, t), t), \quad 0 \leqslant t \leqslant T.$$

From the convexity of each f_i in its first argument, if follows from [8, p. 242] that

$$\sum_{i=1}^{m} f_i(z(t), t) \geqslant \sum_{i=1}^{m} f_i(0, t) + \sum_{k=1}^{m} a_k(t)\, z_k(t),$$

where

$$a_k(t) = \sum_{i=1}^{m} \nabla_k f_i(0, t).$$

Set $\theta_0 = \max\left\{0, \sup_{t\in[0,T]}\left\{-\sum_{i=1}^{m} f_i(0,t)\right\}\right\}$ by (9c) and observe that by assumption (6)

$$A = \inf_t \min_k a_k(t) > 0.$$

Since z is feasible and therefore satisfies constraint (1), it then follows that

$$A\sum_{k=1}^{n} z_k(t) \leqslant \theta_0 + \sum_{i=1}^{m} h_i(y(z,t),t), \quad 0 \leqslant t \leqslant T. \tag{10}$$

Define

$$[\nabla g_j(\eta,s)] = \{\nabla_k g_{ij}(\eta,s)\}_{p\times n}, \text{ for } \eta \in E^n,\ s \in [0,T],\ j = 0, \ldots, r,$$

and

$$[\nabla h(\nu,t)] = \{\nabla_i h_i(\nu,t)\}_{m\times p}, \text{ for } \nu \in E^p,\ t \in [0,T],$$

and set

$$G(z,t,s) = \sum_{j=0}^{r} I_{[0,t-\alpha_j]}(s)\,[\nabla h(y(z,t),t)]g_j(z(s),s)$$

and

$$H(z,t,s) = \sum_{j=0}^{r} I_{[0,t-\alpha_j]}(s)\,[\nabla h(y(z,t),t)][\nabla g_j(z(s),s)]$$

where $I_I(\cdot)$ is the indicator function of the set E.

Since h and g_j are concave in their first arguments it follows from [8] and from (3), (7) and (8) that

$$h(y(z,t),t) \leqslant h(0,t) + \int_0^t G(0,t,s)\,ds + \int_0^t H(0,t,s)z(s)\,ds.$$

By (9a) and (9b) we select $\theta_1 \geqslant 0$ and $\theta_2 > 0$, such that

$$\sup_t\left\{\sum_{i=1}^{m} h_i(0,t) + \sum_{i=1}^{m}\int_0^t G_i(0,t,s)\,ds\right\} \leqslant \theta_1$$

and

$$\sup_t \max_k\left\{\sum_{i=1}^{m} H_{ik}(0,t,s)\right\} \leqslant \theta_2.$$

From (10) we have that $\theta_1^* = (\theta_0 + \theta_1)/A$ and $\theta_2^* = \theta_2/A$ are nonnegative and positive constants, respectively, for which

$$\sum_{k=1}^{n} z_k(t) \leqslant \theta_1^* + \theta_2^*\int_0^t \sum_{k=1}^{n} z_k(s)\,ds, \quad 0 \leqslant t \leqslant T.$$

From Lemma 2 we conclude that

$$\sum_{k=1}^{n} z_k(t) \leqslant \theta_1^* \exp(\theta_2^* t) \leqslant \theta_1^* \exp(\theta_2^* T), \ 0 \leqslant t \leqslant T, \tag{11}$$

and hence the set of feasible solutions for Primal Problem A is uniformly bounded on $[0,T]$.

Since ϕ is concave and differentiable in its first argument throughout $[0,T]$, it follows from (9d), [8] and the uniform boundedness property that, for any feasible solutions z and z^0,

$$V(z) - V(z^0) \leqslant T \sum_{k=1}^{n} \sup_t (z_k(t) - z_k^0(t)) \sup_t \nabla_k \phi(z^0(t),t) < \infty$$

and hence V is bounded above for all feasible z.

Let $\bar{V} = lub\ V(z)$, where the least upper bound (*lub*) is taken over all feasible z. Then there exists a sequence $\{z^d\}$ of feasible solutions such that

$$\lim_{d \to \infty} V(z^d) = \bar{V}.$$

Since $\{z^d\}$ is uniformly bounded, it follows from [10] that there exists a $\bar{z}$ to which a subsequence of $\{z^d\}$ converges weakly in $L_n^2[0,T]$. Denote this weakly convergent subsequence itself by $\{z^d\}$; the application of Lemma 1 to each component of z^d then provides uniform boundedness for $\bar{z}$ except possibly on a set of measure zero where, as will be shown later, it can be assumed to be zero.

Since each component of the composite function $h(y(\cdot,t),t)$ is concave in z, it follows from [8], (3) and the chain rule for differentiation that

$$h(y(z^d,t),t) \leqslant h(y(\bar{z},t),t) + \int_0^t H(\bar{z},t,s)(z^d(s) - \bar{z}(s))ds, \quad 0 \leqslant t \leqslant T.$$

Since each entry of the $m \times n$ matrix $H(\bar{z},t,s)$ is bounded and measurable, it follows that each row $H_i(\bar{z},t,s) \in L_n^\infty[0,T] \subseteq L_n^2[0,T]$ and so, by weak convergence,

$$\int_0^t H(z,t,s)(z^d(s) - z(s))ds \to 0, \text{ as } d \to \infty.$$

Thus, by constraint (2)

$$\lim_{d \to \infty} \sup f(z^d(t),t) \leqslant h(y(\bar{z},t),t), \quad 0 \leqslant t \leqslant T. \tag{12}$$

Define $[\nabla f(\eta,t)] = ((\nabla_k f_i(\eta,t))_{m \times n}$, $\eta \in E^n$, $\eta \geqslant 0$; by the convexity of f

$$f(z^d(t),t) \geqslant f(\bar{z}(t),t) + [\nabla f(\bar{z}(t),t)](z^d(t) - \bar{z}(t)), \quad 0 \leqslant t \leqslant T.$$

Therefore, from (12),

$$f(\bar{z}(t),t) \leqslant h(y(\bar{z},t),t) \tag{13}$$

except on a set of measure zero, since by [8], assumption (5) and Lemma 1 we have

$$\lim_{d \to \infty} \sup [\nabla f(\bar{z}(t),t)](z^d(t) - \bar{z}(t)) \geqslant 0$$

except on such a set.

A second application of Lemma 1 to each component of z^d provides

$$-\bar{z}(t) \leqslant \lim_{d \to \infty} \sup(-z^d(t)) \leqslant 0, \text{ a.e. in } [0,T],$$

and consequently $\bar{z}$ is nonnegative except on a set of measure zero. From this result and expression (13), we observe that $\bar{z}$ can violate the constraints of Primal Problem A on, at most, a set of measure zero in $[0,T]$. We define z to be zero on this set of measure zero, as well as for $t < 0$, and equal to $\bar{z}$ on the complement of this set. The feasibility of z is then established by noting that

$$y(z,t) = y(\bar{z},t), \quad 0 \leqslant t \leqslant T,$$

and that

$$\lim_{d \to \infty} \sup f(z^d(t),t) \geqslant f(0,t), \quad 0 < t < T.$$

by the convexity constraint (1), and assumption (6).

By the concavity and differentiability of ϕ

$$\int_0^T \phi(z^d(t),t)dt \leqslant \int_0^T \phi(\bar{z}(t),t)dt + \int_0^T (z^d(t) - \bar{z}(t))' \nabla\phi(\bar{z}(t),t)dt.$$

Therefore, by weak convergence

$$\begin{aligned}\bar{V} &= \lim_{d\to\infty} \int_0^T \phi(z^d(t),t)dt \\ &\leqslant \int_0^T \phi(\bar{z}(t),t)dt = V(\bar{z}).\end{aligned}$$

By the definition of $\bar{V}$ and the feasibility of $\bar{z}$, $V(\bar{z}) \leqslant \bar{V}$, thus $V(\bar{z}) = \bar{V}$ and $\bar{z}$ is an optimal solution for Primal Problem A. Q.E.D.

4. WEAK DUALITY

Before the dual to Primal Problem A is formally stated, a continuous time Lagrangian function and its Frechet differential will be introduced.

For $u \in L_\infty^r[0,T]$ and $w \in L_\infty^s[0,T]$, define

$$L(u,w) = \int_0^T [\phi(u(t),t) + w'(t)F(u,t)]dt \tag{14}$$

where

$$F(u,t) = h(x(u,t),t) - f(u(t),t), \quad 0 \leqslant t \leqslant T,$$

and let $\delta_1 L(u,w;\gamma)$ denote the Frechet differential [7] with respect to its first argument, evaluated at u with the increment $\gamma \in L_\infty^r[0,T]$. The differentiability of each of the functions involved in L insures that the Frechet differential exists and allows $\delta_1 L(u,w;\gamma)$ to be determined by the simple differentiation

$$\delta_1 L(u,w,\gamma) = \frac{d}{d\alpha} L(u + \alpha\gamma, w)\Big|_{\alpha=0}. \tag{15}$$

The Frechet differential has two additional properties that will be used in the ensuing discussion, namely, the linearity of $\delta_1 L(u,w;\gamma)$ in its increment γ and the continuity of $\delta_1 L(u,w;\gamma)$ in γ under the norm

$$\|\gamma\|_\infty = \max_i \|\gamma_i\|_\infty.$$

Here $\|\cdot\|_\infty$ denotes the essential supremum [9, p. 112].

If $\gamma(t) = 0$ for $t < 0$, then from (14) we have

$$\begin{aligned}\delta_1 L(u,w;\gamma) = \int_0^T &\{[\nabla\phi(u(t),t)]'\gamma(t) \\ &+ \int_0^T w'(t)H(u,t,s)\gamma(s)ds - w'(t)[\nabla f(u(t),t)]\gamma(t)\}dt.\end{aligned} \tag{16}$$

An application of Fubini's theorem [9] to interchange the limits of integration enables us to express (16) as

$$\delta_1 L(u,w,\gamma) = \delta L(u,\gamma) + \int_0^T \gamma'(t)F^*(u,w,t)dt, \tag{17}$$

where

$$\delta L(u,\gamma) = \int_0^T [\nabla\phi(u(t),t)]'\gamma(t)dt$$

and

$$F^*(u,w,t) = \int_t^T H'(u,s,t)w(s)ds - [\nabla f(u(t),t)]'w(t),\ 0 \leqslant t \leqslant T. \tag{18}$$

With this notation the dual of Primal Problem A will be shown to be:

Dual Problem A:

Minimize

$$G(u,w) = L(u,w) - \delta_1 L(u,w;u) \tag{19}$$

subject to the constraints

$$u(t),\ w(t) \geqslant 0,\ 0 \leqslant t \leqslant T, \tag{20}$$

$$F^*(u,w,t) + [\nabla\phi(u(t),t)] \leqslant 0,\ 0 \leqslant t \leqslant T, \tag{21}$$

$$u(t) = 0,\ t < 0 \tag{22}$$

and

$$w(t) = 0,\ t > T. \tag{23}$$

THEOREM 2 (Weak Duality): If z and (u,w) are feasible solutions for Primal and Dual Problems A, respectively, then

$$V(z) \leqslant G(u,w).$$

PROOF: By the concavity of ϕ and $-f$ in their first arguments and the concavity of the composite function $h(y(\cdot,t),t)$ in z it follows that L is concave in its first argument and

$$L(z,w) - L(u,w) \leqslant \delta_1 L(u,w;z-u).$$

Thus,

$$\begin{aligned} V(z) - G(u,w) &= L(z,w) - \int_0^T w'(t)F(z,t)dt \\ &\quad - L(u,w) + \delta_1 L(u,w;u) \\ &\leqslant \delta_1 L(u,w;z-u) + \delta_1 L(u,w;u) \\ &\quad - \int_0^T w'(t)F(z,t)dt \\ &= \delta_1 L(u,w;z) - \int_0^T w'(t)F(z,t)dt \end{aligned}$$

by the linearity of the Frechet differential in its increment. By (17) we have

$$\begin{aligned} \delta_1 L(u,w;z) - \int_0^T w'(t)F(z,t)dt &= \int_0^T z'(t)\{[\nabla\phi(u(t),t)] + F^*(u,w,t)\}dt \\ &\quad - \int_0^T w'(t)F(z,t)dt \end{aligned}$$

which is nonpositive by constraints (1), (2), (20) and (21). Q.E.D.

From Theorem 2 it is observed that if there exist feasible solutions, $\bar{z}$ and $(\hat{u},\hat{w})$, for the primal and dual problems and if the corresponding primal and dual objective function values, $V(\bar{z})$ and $G(\hat{u},\hat{w})$, are equal, then these solutions are optimal for their respective problems.

5. THE CONSTRAINT QUALIFICATION.

The constraint qualification introduced here is motivated by the form of the Kuhn-Tucker constraint qualification presented by Zangwill [11] and also by Property 1 given below. The basic theory surrounding this qualification is established to provide a framework for the theorems of Section 6.

PROPERTY 1: If

$$\delta V(z;\gamma) = \int_0^T \gamma'(t)[\nabla\phi(z(t),t)]dt > 0 \tag{24}$$

where $z, \gamma \in L_n^\infty[0,T]$, then there exists a scalar $\sigma > 0$ such that

$$V(z + \tau\gamma) > V(z), \text{ for } 0 < \tau \leqslant \sigma.$$

PROOF: By (15) and (24)

$$\lim_{\tau \downarrow 0} [V(z + \tau\gamma) - V(z)]/\tau = \delta V(z;\gamma) > 0,$$

thus a positive σ can be chosen which is sufficiently small so that

$$V(z + \tau\gamma) > V(z), \text{ for } 0 < \tau \leqslant \sigma. \qquad \text{Q.E.D.}$$

DEFINITION 1: For each z which is feasible for Primal Problem A, define $D(z)$ to be the set of n-vector functions γ for which

(i) $\gamma \in L_n^\infty[0,T]$

(ii) $\gamma(t) = 0$, for $t < 0$

(iii) there exists a scalar $\sigma > 0$ such that

$$z(t) + \tau\gamma(t) \geqslant 0, \ 0 \leqslant t \leqslant T,$$

and

$$F(z + \tau\gamma, t) \geqslant 0, \ 0 \leqslant t \leqslant T,$$

for

$$0 \leqslant \tau \leqslant \sigma.$$

DEFINITION 2: Define $\bar{D}(z)$ to be the closure of $D(z)$ under the norm $||\cdot||_n^\infty$; that is, if a sequence $\{\gamma^d\} \subset D(z)$ is such that $||\gamma^d - \gamma||_n^\infty \to 0$, as $d \to \infty$, then $\gamma \in \bar{D}(z)$.

Henceforth, the Frechet differential of the mapping $F(\cdot,t)$: $L_n^\infty[0,T] \to E^m$ evaluated at z and with increment γ, will be denoted by $\delta F(z;\gamma)_t$. It should be observed that, for any specified value of $t \in [0,T]$, the existence of $\delta F(z;\gamma)_t$ is ensured by the differentiability of f, g_j, and h and that when $\gamma(t) = 0$ for $t < 0$, we have

$$\delta F(z;\gamma)_t = \int_0^t H(z,t,s)\gamma(s)ds - [\nabla f(z(t),t)]\gamma(t). \tag{25}$$

Similarly, the Frechet differential of a component $F_i(\cdot,t)$ of $F(\cdot,t)$, evaluated at z with increment γ, will be denoted by $\delta F_i(z;\gamma)_t$.

DEFINITION 3: For each z which is feasible for Primal Problem A define $\mathscr{L}(z)$ to be the set of n-vector functions γ for which

(i) $\gamma \in L_n^\infty [0,T]$,

(ii) $\gamma(t) = 0$, for $t < 0$,

(iii) $\gamma_k(t) \geqslant 0$ *a.e.* in $T_{1k}(z)$, $k = 1, \ldots, n$,

(iv) $\delta F_i(z;\gamma)_t \geqslant 0$ *a.e.* in $T_{2i}(z)$, $i = 1, \ldots, m$.

where

$$T_{1k}(z) = \{t \in [0,T]: z_k(t) = 0\}, \quad k = 1, \ldots, n$$

and

$$T_{2i}(z) = \{t \in [0,T]: F_i(z,t) = 0\}, \quad i = 1, \ldots, m.$$

In a comparison of the sets $D(z)$ and $\mathscr{L}(z)$ with their finite-dimensional counterparts presented in Zangwill [11], it is observed that $D(z)$ is analogous to the set of "feasible directions" at z and $\mathscr{L}(z)$ is analogous to that set of directions for which the directional derivatives of each of the active constraints at z are nonnegative.

PROPERTY 2: $\bar{D}(z) \subset \mathscr{L}(z)$.

PROOF: Part 1. Let $\gamma \in D(z)$. Then by Definition 1, there exists a scalar $\sigma > 0$ such that $0 \leqslant \tau \leqslant \sigma$ implies $z(t) + \tau\gamma(t) \geqslant 0$, $0 \leqslant t \leqslant T$. Thus, if $z_k(t) = 0$, then $\gamma_k(t) \geqslant 0$.

Assume that $F_i(z,t) = 0$. If $\delta F_i(z;\gamma)_t < 0$, then by the same technique used in the proof of Property 1, it follows that for τ sufficiently small,

$$F_i(z + \tau\gamma, t) < F_i(z,t) = 0.$$

This result contradicts the assumption that $\gamma \in D(z)$ and therefore we conclude that $D(z) \subset \mathscr{L}(z)$.

Part 2. Assume that there is a $\gamma \in L_n^\infty [0,T]$ and a sequence $\{\gamma^d\} \subset D(z)$ such that $\max_k ||\gamma_k^d - \gamma_k||^\infty \to 0$, as $d \to \infty$. Then for all t such that $z_k(t) = 0$, $\gamma_k^d(t) \geqslant 0$, $d = 1, 2, \ldots$. It then follows from convergence in $L^\infty [0,T]$ that $\gamma_k(t) \geqslant 0$ *a.e.* on $T_{1k}(z)$, $k = 1, \ldots, n$.

Assume there exists an i and a set E of positive measure over which $F_i(z,t) = 0$ and $\delta F_i(z;\gamma)_t < 0$ for all $t \in E$. By the continuity of $\delta F_i(z;\cdot)_t$ in the L^∞ norm [7], we can choose a d^* sufficiently large such that for $d \geqslant d^*$

$$\delta F_i(z;\gamma^d)_t < 0$$

over some subset of E which has positive measure. This result yields a contradiction to Part 1 since it was assumed $\{\gamma^d\} \subset D(z)$ and we can therefore conclude that $\bar{D}(z) \subset \mathscr{L}(z)$. Q.E.D.

DEFINITION 4 (Constraint Qualification): Primal Problem A will be said to satisfy the Constraint Qualification if the problem is feasible and if

$$\bar{D}(\bar{z}) = \mathscr{L}(\bar{z}),$$

where $\bar{z}$ is an optimal solution to the problem.

In more general problems where convexity and concavity properties are not assumed, the purpose of the Constraint Qualification would be to eliminate "cusps" in the feasible region. For example, the constraints

$$z_1(t) \geqslant 0,\ z_2(t) \geqslant 0,\ 0 \leqslant t \leqslant T,$$

and

$$[1 - z_1(t)]^3 - z_2(t) \geqslant 0,\ 0 \leqslant t \leqslant T,$$

do not satisfy the Constraint Qualification when $\bar{z}(t) \equiv (1,0),\ 0 \leqslant t \leqslant T$, since $(1/2, 0) \in \mathscr{D}(\bar{z})$ but $(1/2, 0) \notin \bar{D}(\bar{z})$.

In problems such as Primal Problem A where convexity and concavity properties are assumed, violations of the Constraint Qualification can be shown to arise when the constraints take the form of equalities on some set of positive measure. For example, consider the constraints

$$z_1(t) \geqslant 0,\ z_2(t) \geqslant 0,\ 0 \leqslant t \leqslant T,$$

and

$$[z_1(t) + z_2(t) - 1]^2 \leqslant 1 - I_E(t),\ 0 \leqslant t \leqslant T,$$

where E is a set of positive measure in [0,T] and $I_E(\cdot)$ is its indicator function. It is observed that for $\bar{z}(t) \equiv (1/2, 1/2)$, we have $(1,1) \in \mathscr{D}(\bar{z})$ but $(1,1) \notin \bar{D}(z)$, thus the Constraint Qualification is not satisfied.

THEOREM 3: If $\bar{z}$ is optimal for Primal Problem A, then under the Constraint Qualification

$$\delta V(\bar{z};\gamma) \leqslant 0, \text{ for all } \gamma \in \mathscr{D}(\bar{z}).$$

PROOF: Part 1. Suppose there exists a $\gamma \in D(\bar{z})$ such that $\delta V(\bar{z};\gamma) > 0$. Then by Property 1 there exists a $\sigma > 0$ such that $0 < \tau \leqslant \sigma$ implies $V(\bar{z} + \tau\gamma) > V(\bar{z})$; however, since $\gamma \in D(\bar{z})$ we can choose a σ_0 sufficiently small so that $\bar{z} + \sigma_0\gamma$ is feasible for Primal Problem A. Thus, by contradiction of the optimality of $\bar{z}$, we can conclude that $\delta V(\bar{z};\gamma) \leqslant 0$, for all $\gamma \in D(\bar{z})$.

Part 2. Let $\{\gamma^d\}$ be a sequence of functions in $D(\bar{z})$ and let γ^0 be such that $\max_k ||\gamma_k^d - \gamma_k^0||^\infty \to 0$, as $d \to \infty$. It then follows from Part 1 and the continuity of $\delta V(\bar{z};\cdot)$ that

$$\delta V(\bar{z};\gamma^0) = \lim \delta V(\bar{z};\gamma^d) \leqslant 0.$$

Thus, $\delta V(\bar{z};\gamma) \leqslant 0$ for all $\gamma \in \bar{D}(\bar{z})$. Q.E.D.

6. DUALITY AND RELATED THEOREMS

In proving strong duality and its related theorems two additional assumptions will be made. These are:

(26) $$H(\bar{z},t,s) \geqslant 0,\ 0 \leqslant s \leqslant t \leqslant T$$

and

$$F(\bar{z},t) - \delta F(\bar{z};\bar{z})_t \geqslant 0, \quad 0 \leqslant t \leqslant T, \tag{27}$$

where $\bar{z}$ is an optimal solution for Primal Problem A. It will be shown in Corollary 1 that assumption (26) is implied if $z(t) \equiv 0$ is feasible.

THEOREM 4 (Strong Duality): Under the Constraint Qualification and assumptions (26) and (27), there exists an optimal solution $(\bar{u},\bar{w})$ for Dual Problem A such that $\bar{u} = \bar{z}$ and $G(\bar{z},\bar{w}) = V(\bar{z})$.

Before proving Theorem 4 the following linearized problem, called *Primal Problem A'*, will be considered:

Maximize

$$\delta V(\bar{z};z - \bar{z})$$

subject to the constraints

$$z(t) \geqslant 0, \ 0 \leqslant t \leqslant T, \tag{28}$$

$$F(\bar{z},t) + \delta F(\bar{z};z - \bar{z})_t \geqslant 0, \ 0 \leqslant t \leqslant T, \tag{29}$$

and

$$z(t) = 0, \text{ for } t < 0. \tag{30}$$

LEMMA 3: Under the Constraint Qualification, $\bar{z}$ is an optimal solution for Primal Problem A'.

PROOF: If $\hat{z}$ is feasible for Primal Problem A', then

$$\hat{z}(t) - \bar{z}(t) \geqslant 0, \text{ for } t \in T_{1k}(\bar{z}), \ k = 1, \ldots, n,$$

and

$$\delta F_i(\bar{z},\hat{z} - \bar{z})_t \geqslant 0, \text{ for } t \in T_{2i}(\bar{z}), \ i = 1, \ldots, m,$$

and therefore $(\hat{z} - \bar{z}) \in \mathscr{D}(\bar{z})$. It then follows from Theorem 3 that, under the Constraint Qualification,

$$\delta V(\bar{z};\hat{z} - \bar{z}) \leqslant 0$$

for all $\hat{z}$ satisfying (28), (29) and (30). The optimality of $\bar{z}$ follows since $\bar{z}$ is feasible for Primal Problem A' and since $\delta V(\bar{z};0) = 0$. Q.E.D.

PROOF OF THEOREM 4: We rewrite Primal Problem A' in the form

maximize

$$\int_0^T a'(t)z(t)dt$$

subject to the constraints

$$z(t) \geqslant 0, \ 0 \leqslant t \leqslant T,$$

and

$$B(t)z(t) \leqslant c(t) + \int_0^t K(t,s)z(s)ds, \ 0 \leqslant t \leqslant T,$$

where $a(t) = [\nabla\phi(\bar{z}(t),t)]$, $B(t) = [\nabla f(\bar{z}(t),t)]$, $c(t) = F(\bar{z},t) - \delta F(\bar{z};\bar{z})_t$, and $K(t,s) = H(\bar{z},t,s)$. From assumptions (5), (6), (26) and (27) it is observed that the matrices $a(t)$, $B(t)$, $c(t)$ and $K(t,s)$ satisfy the requirements of Grinold's Duality Theorem [3]. Therefore, there exists an m-vector function $\bar{w}$ satisfying

(31) $$\bar{w}(t) \geqslant 0, \ 0 \leqslant t \leqslant T,$$

and

(32) $$B'(t)\bar{w}(t) \geqslant a(t) + \int_t^T K'(s,t)\bar{w}(s)ds, \ 0 \leqslant t \leqslant T,$$

such that

(33) $$\int_0^T \bar{w}'(t)\ c(t)dt = \int_0^T a'(t)\bar{z}(t)dt.$$

Setting $\bar{w}(t) = 0$ for $t > T$, we observe from the identities (14), (17), and (18) that expressions (32) and (33) can be expressed as

(32′) $$F^*(\bar{z},\bar{w},t) + [\nabla\phi(\bar{z}(t),t)] \leqslant 0, \ 0 \leqslant t \leqslant T,$$

and

(33′) $$L(\bar{z},\bar{w}) - \delta_1 L(\bar{z},\bar{w};\bar{z}) = V(\bar{z}),$$

respectively. From (31) and (32′) and the fact that $\bar{w}(t) = 0$ for $t > T$, it then follows that $(\bar{z},\bar{w})$ is feasible for Dual Problem A and, from (19) and (33′)

(34) $$G(\bar{z},\bar{w}) = V(\bar{z}).$$

Finally, by the weak duality established in Theorem 2, it is concluded from (34) that $(\bar{z},\bar{w})$ is an optimal solution for Dual Problem A. Q.E.D

In order to apply Theorem 4 in practice, it is desirable to be able to verify conditions (26) and (27) without prior knowledge of the optimal solution $\bar{z}$. The following corollary provides this capability.

COROLLARY 1: If

$$\nabla_k g_{ji}(\eta,t)/\partial\eta_k = \partial g_{ji}(\eta,t)/\partial\eta_k \geqslant 0,$$

(35) $$j = 0, \ldots, r, \ i = 1, \ldots, p, \ k = 1, \ldots, n,$$

for $\eta \in E^n$, $\eta \geqslant 0$, and $0 \leqslant t \leqslant T$,

(36) $$F(0,t) \geqslant 0, \ \ 0 \leqslant t \leqslant T,$$

then under the Constraint Qualification there exists an optimal solution $(\bar{u},\bar{w})$ for Dual Problem A such that $\bar{u} = \bar{z}$ and $G(\bar{z},\bar{w}) = V(\bar{z})$.

PROOF: We have from (8) and (35) that

$$H(\bar{z},t,s) = \sum_{j=0}^{r} I_{[0,t-\alpha_j]}(s)\,[\nabla h(\gamma(\bar{z},t),t)]\,[\nabla g_j(\bar{z}(s),s)] \geqslant 0, \ 0 \leqslant t \leqslant T,$$

and by (36) and the concavity of F that

$$F(\bar{z},t) - \delta F(\bar{z};\bar{z})_t \geqslant F(0,t) \geqslant 0, \ \ 0 \leqslant t \leqslant T.$$

From these results it follows that the conditions of Theorem 4 are satisfied. Q.E.D.

THEOREM 5 (Complementary Slackness Principle): If $\bar{z}$ and $(\bar{z},\bar{w})$ are optimal solutions for the Primal and Dual Problems A, then

$$\int_0^T \bar{w}'(t)F(\bar{z},t)dt = 0 \tag{37}$$

and

$$\int_0^T \bar{z}'(t)\{F^*(\bar{z},\bar{w},t) + [\nabla\phi(\bar{z}(t),t)]\}dt = 0. \tag{38}$$

PROOF: Since $\bar{z}(t) \geqslant 0$ and $F^*(\bar{z},\bar{w},t) + [\nabla\phi(\bar{z}(t),t)] \leqslant 0$, $0 \leqslant t \leqslant T$, it follows from identity (17) that

$$\int_0^T \bar{z}'(t)\{F^*(\bar{z},\bar{w},t) + [\nabla\phi(\bar{z}(t),t)]\}dt = \delta_1 L(\bar{z},\bar{w},\bar{z}) \leqslant 0,$$

and therefore, by (33′)

$$L(\bar{z},\bar{w}) - V(\bar{z}) = \int_0^T \bar{w}'(t)F(\bar{z},t)dt \leqslant 0$$

Since $\bar{w}(t) \geqslant 0$ and $F(\bar{z},t) \geqslant 0$, $0 \leqslant t \leqslant T$, it also follows that

$$\int_0^T \bar{w}'(t)F(\bar{z},t)dt \geqslant 0, \tag{39}$$

thus the equality in (37) is established.

Similarly, (33′) and (39) imply that

$$\delta_1 L(\bar{z},\bar{w},\bar{z}) \geqslant 0$$

and therefore, by (17)

$$\int_0^T \bar{z}'(t)\{F^*(\bar{z},\bar{w},t) + [\nabla\phi(\bar{z}(t),t)]\}dt \geqslant 0.$$

Since $\bar{z}(t) \geqslant 0$ and $F^*(\bar{z},\bar{w},t) + [\nabla\phi(\bar{z}(t),t)] \leqslant 0$, $0 \leqslant t \leqslant T$, we have

$$\int_0^T \bar{z}'(t)\{F^*(\bar{z},\bar{w},t) + [\nabla\phi(\bar{z}(t),t)]\}dt \leqslant 0$$

and thus the equality in (38) is established. Q.E.D.

THEOREM 6 (Kuhn-Tucker Conditions): Assume that (35) and (36) are satisfied for Primal Problem A. Then under the Constraint Qualification $\bar{z}$ is an optimal solution if and only if there exists an m-vector function $\hat{w}$ such that

(i) $F^*(\bar{z},\hat{w},t) + [\nabla\phi(\bar{z}(t),t)] \leqslant 0$, $0 \leqslant t \leqslant T$,

(ii) $\int_0^T \bar{z}'(t)\{F^*(\bar{z},\hat{w},t) + [\nabla\phi(\bar{z}(t),t)]\}dt = 0$

(iii) $\int_0^T \hat{w}'(t)F(\bar{z},t)dt = 0$

(iv) $\hat{w}(t) \geqslant 0$, $0 \leqslant t \leqslant T$ and $\hat{w}(t) = 0$, $t > T$

PROOF:

Necessity: The necessity of the conditions follows from [illegible] and [illegible] since the m-vector function $\hat{w}$ of the optimal solution $(\bar{z},\hat{w})$ [illegible] satisfies conditions (i) through (iv).

Sufficiency: Let z be feasible for Primal Problem A. Then since V is concave

$$V(z) - V(\bar{z}) \leqslant \delta V(\bar{z}; z - \bar{z})$$
$$= \int_0^T [z(t) - \bar{z}(t)]'[\nabla \phi(\bar{z}(t), t)]\, dt.$$

Since $z(t) \geqslant 0$, $0 \leqslant t \leqslant T$, it follows from conditions (i) and (ii) that

$$V(z) - V(\bar{z}) \leqslant -\int_0^T [z(t) - \bar{z}(t)]' F^*(\bar{z}, \hat{w}, t)\, dt,$$

and by (18), (25) and Fubini's Theorem [9]

$$\int_0^T [z(t) - \bar{z}(t)]' F^*(\bar{z}, \hat{w}, t)\, dt = \int_0^T \hat{w}'(t) \delta F(\bar{z}; z - \bar{z}),\, dt.$$

By (i), (iii) and the concavity of F,

$$-\int_0^T \hat{w}'(t) \delta F(\bar{z}; z - \bar{z}),\, dt \leqslant -\int_0^T \hat{w}'(t)[F(z,t) - F(\bar{z},t)]\, dt$$
$$= -\int_0^T \hat{w}'(t) F(z,t)\, dt$$

which is nonpositive since $\hat{w}(t) \geqslant 0$ and $F(z,t) \geqslant 0$, $0 \leqslant t \leqslant T$. Thus, $V(z) \leqslant V(\bar{z})$ and $\bar{z}$ is an optimal solution for Primal Problem A. Q.E.D.

7. EXAMPLE — WATER STORAGE PROBLEM

In the water storage problem posed in [4], the hydroelectric company incurred a penalty if it could not meet a prescribed demand for power. This penalty was characterized in the objective function

$$\int_0^T \psi(D(t) - P(t))\, dt$$

where $[0,T]$ represents a planning period of specified duration, $D(t)$ is the demand rate, $P(t)$ is the production rate of hydroelectric power, and ψ is the penalty function which was assumed to be strictly convex. The imposition of such a penalty favors the consumer or a middleman utility company which retails electric power to the consumers. In short, it characterizes a "buyers market."

If there is, in fact, a pending energy crisis, it seems appropriate to consider a "sellers market" where the demand for power exceeds production capacity and a premium is paid to the hydroelectric company for any power which it produces beyond some prescribed level. In the case where the hydroelectric company is supplying power directly to the consumer, these premiums may take the form of increasing prices per unit beyond some allotment level. When the hydroelectric company is supplying a middleman, the premiums may represent an incentive policy which encourages maximum production during peak demand periods.

The premiums to the hydroelectric company will be represented by

$$\int_0^T \pi\,(P(t) - A(t))\, dt$$

where $[0,T]$ represents the planning period, $P(t)$ is the power production rate, $A(t)$ is the prescribed aggregate allotment or incentive level, and π is the premium function which is assumed to be differentiable and concave with a positive slope at zero.

For the dynamics of the problem, we assume a confluent system of rivers supplying water to a hydroelectric plant on the main stream with r of its tributaries also having their own

hydroelectric plants. The variables and parameters which relate to the dam, reservoir and plant on the main stream will be subscripted by 0, and those for the r dammed tributaries by j, $j = 1, \ldots, r$.

We let Ω_j denote the initial store of water in reservoir j and θ_j the capacity of reservoir j. The rate of spillage and rate of discharge through the turbines of dam j at time t are denoted by $s_j(t)$ and $d_j(t)$, respectively. The rates of inflow of water into the reservoirs on the dammed tributaries are $\xi_j(t)$ $j = 1, \ldots, r$, and that into the main reservoir from its undammed tributaries is $\xi_0(t)$.

It is assumed that it takes α_j, $j = 1, \ldots, r$ units of time for the water released from dam j to reach the main reservoir and that there is no spillage or discharge through the dams on the tributaries for at least α units of time prior to the start of the planning period, where $\alpha = \max\{\alpha_j\}$. The store of water in reservoir j at time t can then be expressed as

$$W_j(t) = \Omega_j + \int_0^t (\xi_j(t') - s_j(t') - d_j(t'))\,dt'$$

for $j = 1, \ldots, r$, and

$$W_0(t) = \Omega_0 + \int_0^t \left\{\xi_0(t') - s_0(t') - d_0(t') + \sum_{j=1}^{r} (s_j(t' - \alpha_j) + d_j(t' - \alpha_j))\right\} dt'$$

for the main reservoir.

The power production rate for a given rate of discharge d is assumed to be proportional to d. In [4], it was necessary to assume the factor of proportionality to be unity. Here we allow this factor to be proportional to the head of water in the reservoir, an assumption which is consistent with constant turbine efficiency. The head is the difference h between the surface level of the reservoir and the tailwaters below the dam and is therefore dependent primarily upon the store of water W in the reservoir.

The relationship between $h_j(t)$, the head of reservoir j, and $W_j(t)$ will be represented by $h_j(t) = h_j^*(W_j(t))$, where h_j^* is an increasing concave differentiable function. The functions h_j^* owe their concavity to the shapes of the reservoirs which are assumed to yield a continuously disproportionate increase in reservoir surface area as the store of water increases. The production rate for the jth hydroelectric plant is then expressible as

$$p_j(t) = d_j(t) \circ h_j^*(W_j(t)),$$

in which case the production rate for the entire system becomes

$$P(t) = \sum_{j=0}^{r} p_j(t).$$

Assuming the role of the hydroelectric company, we want to select our water storage policy (s,d) so as to maximize the premium payments over the planning period. This problem takes the form

maximize

$$\pi(s,d) = \int_0^T \pi(P(t) - A(t))\,dt$$

subject to

$$0 \leqslant s_j(t) \leqslant \beta_j(t)$$

$$0 \leqslant d_j(t) \leqslant \phi_j$$

$$0 \leqslant W_j(t) \leqslant \theta_j$$

$j = 0, \ldots, r$, and

$$A(t) \leqslant P(t)$$

for $0 \leqslant t \leqslant T$, where $\beta_j(t)$ is the maximum allowable spillage rate through dam j and ϕ_j is the turbine capacity of plant j.

Through proper association of the terms of this model with those of Primal Problem A it can be shown through application of Theorem 1 that feasibility ensures the existence of an optimal water storage policy which will maximize the total premium payment.

REFERENCES

[1] Farr, W.H. and M.A. Hanson, "Continuous Time Programming with Nonlinear Constraints," Journal of Mathematical Analysis and Applications *45*, 96-115 (1974).

[2] Farr, W.H. and M.A. Hanson, "Continuous Time Programming with Nonlinear Time-Delayed Constraints," Journal of Mathematical Analysis and Applications *46*, 41-60 (1974).

[3] Grinold, R., "Continuous Programming Part One: Linear Objectives," Journal of Mathematical Analysis and Applications *28*, 32-51 (1969).

[4] Koopmans, T.C., "Water Storage in a Simplified Hydroelectric System," *Proceedings of the First International Conference on Operational Research*, M. Davies, R.T. Eddison and T. Page, Editors (Operations Research Society of America, Baltimore, 1957).

[5] Kuhn, H.W. and A.W. Tucker, "Nonlinear Programming," *Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probabilities*, 481-492, J. Neyman, Editor (University of California Press, Berkeley, 1951).

[6] Levinson, N., "A Class of Continuous Linear Programming Problems," Journal of Mathematical Analysis and Applications *16*, 73-83 (1966).

[7] Luenberger, D.G., *Optimization by Vector Space Methods* (Wiley, New York, N.Y., 1969).

[8] Rockafellar, R.T., *Convex Analysis* (Princeton University, Princeton, New Jersey, 1970).

[9] Royden, H.L., *Real Analysis* (MacMillan, New York, 1968).

[10] Taylor, A.E., *Introduction to Functional Analysis* (Wiley, New York, N.Y., 1958).

[11] Zangwill, W.I., *Nonlinear Programming: A Unified Approach* (Prentice Hall, Englewood Cliffs, New Jersey, 1969).

EQUALITIES IN TRANSPORTATION PROBLEMS AND CHARACTERIZATIONS OF OPTIMAL SOLUTIONS*

Kenneth O. Kortanek

Department of Mathematics,
Carnegie-Mellon University
Pittsburgh, Pennsylvania

Maretsugu Yamasaki

Department of Mathematics
Shimane University
Matsue, Shimane, Japan

ABSTRACT

This paper considers the classical finite linear transportation Problem (I) and two relaxations, (II) and (III), of it based on papers by Kantorovich and Rubinstein, and Kretschmer. Pseudo-metric type conditions on the cost matrix are given under which Problems (I) and (II) have common optimal value, and a proper subset of these conditions is sufficient for Problems (II) and (III) to have common optimal value. The relationships between the three problems provide a proof of Kantorovich's original characterization of optimal solutions to the standard transportation problem having as many origins as destinations. The results are extended to problems having cost matrices which are nonnegative row-column equivalent

1. INTRODUCTION WITH PROBLEM SETTING

Over 25 years ago Kantorovich in his classic paper, "On the translocation of masses" [4], formulated generalized transportation problems which are continuous analogs of the well-known transportation problem in the theory of finite linear programming. He raised the question of characterizing optimal solutions to those problems whose finite dimensional versions have the same number of origins as destinations. As is well known, optimal solutions to the standard finite dimensional transportation problem having "m" origins and "n" destinations are characterized by means of a system of linear inequalities involving m row numbers and n column numbers which together comprise a feasible list of dual variables.

Within the finite dimensional context $m = n$, Kantorovich's goal was to use only n numbers in a linear inequality system characterization of an optimal solution rather than the standard $2n$ (row plus column) numbers. In order to accomplish this, three conditions defining a pseudo-metric were imposed on the cost coefficient matrix. Actually, the triangle inequality condition on unit costs is what Gomory and Hu later termed "reasonable costs" in their network

*The work of the first author was supported in part by National Science Foundation Grants ENG76-05191 and ENG78-25488

studies [3], Section 2. Violation of this particular condition is also related to the "more for less" paradox in the transportation model, see Ryan [7].

The original application of the pseudo-metric conditions involved subtleties which were later clarified in Kantorovich-Rubinstein [5] but for a transformed version of the standard transportation problem, which we state as Problem III in the next section. In attempting to give a proof of Kantorovich's characterization, Kretschmer [6] introduced yet another transformation of the standard problem, which we shall term Problem II in the next section.

The basic purpose of this paper is to delineate the key relationships between these three problems: the standard transportation Problem I, the Kretschmer transformed Problem II, and the Kantorovich-Rubinstein Problem III. The results we obtain depend on how the three pseudo-metric cost conditions, denoted (C.1) through (C.3) in Sections 3 and 4, are coupled together.

Our main application is to obtain a proof of the originally sought for characterization of optimal solutions of the standard transportation problem where the number of origins equals the number of destinations. We are not prepared at this time however to state that we have industrial or public sector applications of the type II or type III transportation models.

2. THE KANTOROVICH-RUBINSTEIN AND KRETSCHMER TRANSFORMS OF THE STANDARD TRANSPORTATION PROBLEM

Let c_{ij}, a_i and $b_j (i = 1, \ldots, n; j = 1, \ldots, n)$ be nonnegative real numbers and assume that a_i and b_j satisfy

$$\sum_{i=1}^{n} a_i = \sum_{j=1}^{n} b_j > 0. \tag{1.1}$$

The original transportation problem may be expressed as follows:

(I) Determine the minimum value M of

$$\sum_{i=1}^{n} \sum_{j=1}^{n} c_{ij} x_{ij} \tag{1.2}$$

subject to the condition that x_{ij} are nonnegative and

$$\sum_{j=1}^{n} x_{ij} = a_i \quad (i = 1, \ldots, n), \tag{1.3}$$

$$\sum_{i=1}^{n} x_{ij} = b_j \quad (j = 1, \ldots, n).$$

Let us consider the following transportation problems which were studied in [5] and [6]:

(II) Determine the minimum value N of

$$\sum_{i=1}^{n} \sum_{j=1}^{n} c_{ij} (x_{ij} + y_{ij}) \tag{1.4}$$

subject to the condition that x_{ij} and y_{ij} are nonnegative and

$$\sum_{j=1}^{n} (x_{ij} - y_{ij}) = a_i \quad (i = 1, \ldots, n), \tag{1.5}$$

$$\sum_{i=1}^{n} (x_{ij} - y_{ij}) = b_j \quad (j = 1, \ldots, n).$$

(III) Determine the minimum value V of

$$\sum_{i=1}^{n}\sum_{j=1}^{n} c_{ij} z_{ij} \tag{1.6}$$

subject to the condition that z_{ij} are nonnegative and

$$\sum_{j=1}^{n} z_{ij} - \sum_{j=1}^{n} z_{ji} = a_i - b_i \quad (i = 1, \ldots, n). \tag{1.7}$$

Program I of course is the classical transportation problem which may be solved by the well-known row and column number method ([1],[2]) and other more modern, large scale programming methods. The row and column number method easily extends to solving Program II. On the other hand, the structural matrix of Program III is a network incidence matrix, and so III is an uncapacitated network problem.

It is clear that $V \leqslant M$ and $N \leqslant M$ and in this sense Problems II and III are relaxations of Problem I. We shall study when one of the equalities $V = N$, $V = M$, and $M = N$ holds.

3. THE EQUALITY N = V OF PROBLEMS II AND III

First we have

LEMMA 1: The inequality $V \leqq N$ holds if the following condition is fulfilled:

$$c_{ij} = c_{ji} \text{ for all } i \text{ and } j. \tag{C.1}$$

PROOF: There exists an optimal solution x_{ij} and y_{ij} of Problem (II), i.e., x_{ij} and y_{ij} are nonnegative and satisfy (1.5) and

$$N = \sum_{i=1}^{n}\sum_{j=1}^{n} c_{ij}(x_{ij} + y_{ij}).$$

Then

$$\left(\sum_{j=1}^{n} x_{ij} + \sum_{j=1}^{n} y_{ji}\right) - \left(\sum_{j=1}^{n} x_{ji} + \sum_{j=1}^{n} y_{ij}\right) = a_i - b_i.$$

Taking $z_{ij} = x_{ij} + y_{ji}$, we see that z_{ij} are nonnegative and satisfy (1.6), so that by condition (C.1)

$$V \leqq \sum_{i=1}^{n}\sum_{j=1}^{n} c_{ij} z_{ij} = N.$$

THEOREM 1: The equality $V = N$ holds if condition (C.1) and the following condition are fulfilled:

$$c_{ii} = 0 \text{ for all } i. \tag{C.2}$$

PROOF: There exists an optimal solution z_{ij} of Problem (III), i.e., z_{ij} are nonnegative and satisfy (1.7) and

$$V = \sum_{i=1}^{n}\sum_{j=1}^{n} c_{ij} z_{ij}.$$

Then

$$\sum_{j=1}^{n} z_{ij} + b_i = \sum_{j=1}^{n} z_{ji} + a_i = d_i \geqq 0.$$

Let us take $x_{ij} = 0$ if $i \neq j$ and $x_{ii} = d_i$ and put $y_{ij} = z_{ji}$. Then x_{ij} and y_{ij} are nonnegative and satisfy (1.5), so that

$$N \leqq \sum_{i=1}^{n} \sum_{j=1}^{n} c_{ij}(x_{ij} + y_{ij}) = \sum_{i=1}^{n} \sum_{j=1}^{n} c_{ij} z_{ji} = V$$

by conditions (C.1) and (C.2).

We show by an example that the equality $N = V$ does not hold in general if we omit condition (C.2).

EXAMPLE 1: Let $n = 2$ and take

$$c_{11} = c_{22} = 1, \quad c_{12} = c_{21} = 2,$$
$$a_1 = 1, \quad a_2 = 2, \quad b_1 = 2, \quad b_1 = 1.$$

Then we easily see that $V = 2$ and $M = N = 4$.

4. THE EQUALITY M = N OF PROBLEMS I AND II

Next we show that the equality $M = N$ does not hold in general even if both conditions (C.1) and (C.2) are fulfilled.

EXAMPLE 2: Let $n = 3$ and take

$$c_{11} = c_{22} = c_{33} = 0, \quad c_{12} = c_{21} = 20,$$
$$c_{13} = c_{31} = c_{23} = c_{32} = 1,$$
$$a_1 = 3/2, \quad a_2 = 1/2, \quad a_3 = 1/4,$$
$$b_1 = b_2 = 1, \quad b_3 = 1/4.$$

By special methods of linear programming (see, for instance [1]), we see that $M = \frac{11}{2}$ and an optimal solution of Problem (I) is given by $x_{11} = 1$, $x_{22} = 1/2$ $x_{12} = x_{13} = x_{32} = 1/4$ and $x_{21} = x_{31} = x_{23} = x_{33} = 0$. We have $N = 1$. An optimal solution of Problem (II) is given by $x_{11} = 1$, $x_{22} = 1/2$, $x_{13} = x_{32} = 1/2$, $x_{12} = x_{21} = x_{23} = x_{31} = x_{33} = 0$, $y_{33} = 1/4$ and $y_{ij} = 0$ if $(i, j) \neq (3, 3)$.

Our main result is the following one.

THEOREM 2: The equality $M = N$ holds if the following condition is fulfilled:

(C.3) $\qquad c_{ij} \leqslant c_{iq} + c_{pj} + c_{pq}$ for all i, j, p, q.

PROOF: There exists an optimal solution x_{ij} and y_{ij} of Problem (II). In case $z_{ij} = x_{ij} - y_{ij}$ is nonnegative for each i, j, we see that z_{ij} is a feasible solution of Problem (I), so that

$$M \leqq \sum_{i=1}^{n} \sum_{j=1}^{n} c_{ij} z_{ij} \leqq \sum_{i=1}^{n} \sum_{j=1}^{n} c_{ij}(x_{ij} + y_{ij}) = N.$$

We consider the case where some $x_{ij} - y_{ij}$ are negative. We may assume that $\min(x_{ij}, y_{ij}) = 0$ for all i and j. There exist p and q such that $0 = x_{pq} < y_{pq}$. Then we have by (1.5)

$$\sum_{i=1}^{n} x_{iq} \geqq y_{pq} \text{ and } \sum_{j=1}^{n} x_{pj} \geqq y_{pq}.$$

Let us define A_i, B_j and d_{ij} by

$$A_p = B_q = 0,$$

$$A_i = x_{iq}y_{pq} / \sum_{i=1}^{n} x_{iq} \quad (i \neq p), \quad B_j = x_{pj}y_{pq} / \sum_{j=1}^{n} x_{pj} \quad (j \neq q),$$

$$d_{ij} = A_i B_j / y_{pq}.$$

Then

$$\sum_{j=1}^{n} d_{ij} = A_i \leqq x_{iq}, \quad \sum_{i=1}^{n} d_{ij} = B_j \leqq x_{pj}, \tag{4.1}$$

$$\sum_{i=1}^{n} A_i = \sum_{j=1}^{n} B_j = y_{pq}. \tag{4.2}$$

We define x'_{ij} and y'_{ij} by

$$\begin{aligned} x'_{ij} &= x_{ij} + d_{ij} && \text{if } i \neq p \text{ and } j \neq q, \\ x'_{pj} &= x_{pj} - B_j && \text{if } j \neq q, \\ x'_{iq} &= x_{iq} - A_i && \text{if } i \neq p, \\ y'_{ij} &= y_{ij} && \text{if } i \neq p \text{ or } j \neq q, \\ x'_{pq} &= y'_{pq} = 0. \end{aligned} \tag{4.3}$$

Then x'_{ij} and y'_{ij} are nonnegative and satisfy (1.5) and

$$\begin{aligned} N &\leqq \sum_{i=1}^{n} \sum_{j=1}^{n} c_{ij}(x'_{ij} + y'_{ij}) \\ &= \sum_{i=1}^{n} \sum_{j=1}^{n} c_{ij}(x_{ij} + y_{ij}) + \sum_{i=1}^{n} \sum_{j=1}^{n} d_{ij}[c_{ij} - c_{iq} - c_{pq} - c_{pj}] \\ &\leqq \sum_{i=1}^{n} \sum_{j=1}^{n} c_{ij}(x_{ij} + y_{ij}) = N \end{aligned}$$

by condition (C.3). Repeating the above procedure (4.3) a finite number of times,† we obtain x^*_{ij} which are nonnegative and satisfy (1.3) and

$$M \leqq \sum_{i=1}^{n} \sum_{j=1}^{n} c_{ij}x^*_{ij} \leqq \sum_{i=1}^{n} \sum_{j=1}^{n} c_{ij}(x_{ij} + y_{ij}) = N.$$

Hence, $M = N$.

THEOREM 3: Let k_{ij}, f_i and g_j be nonnegative numbers and assume that condition (C.3) holds for k_{ij} instead of c_{ij}. If $c_{ij} = k_{ij} + f_i + g_j$, then $M = N$.

†This number is at most the number of $\{y_{ij} > 0\}$.

PROOF: Denote by $M(k)$ and $N(k)$ the values of Problems (I) and (II) respectively if c_{ij} are replaced by k_{ij}. Then we have $M = M(k) + C_{fg}$ with

$$C_{fg} = \sum_{i=1}^{n} f_i a_i + \sum_{j=1}^{n} g_j b_j.$$

Let x_{ij} and y_{ij} be nonnegative and satisfy (1.5). Then

$$\sum_{i=1}^{n}\sum_{j=1}^{n} c_{ij}(x_{ij} + y_{ij}) = \sum_{i=1}^{n}\sum_{j=1}^{n} k_{ij}(x_{ij} + y_{ij}) + \sum_{i=1}^{n} f_i\left[a_i + \sum_{j=1}^{n} y_{ij}\right] + \sum_{j=1}^{n} g_j\left[b_j + \sum_{i=1}^{n} y_{ij}\right]$$

$$\geqq \sum_{i=1}^{n}\sum_{j=1}^{n} k_{ij}(x_{ij} + y_{ij}) + C_{fg} \geqq N(k) + C_{fg}.$$

Thus, $N \geqq N(k) + C_{fg}$. Since $N(k) = M(k)$ by Theorem 2, we have $N \geqq M(k) + C_{fg} = M$, and hence $M = N$.

As is well known, two transportation problems of type (I) with costs $\{c_{ij}\}$ and $\{c_{ij} + f_i + g_j\}$ respectively, are equivalent for any list of real numbers $\{f_i\}$, $\{g_j\}$, $i = 1, \ldots, m$; $j = 1, \ldots, n$. The following example shows that the nonnegativity of all the f_i and g_j is required in Theorem 3.

EXAMPLE 3: Let $n = 2$ and take $k_{11} = 0$, $k_{12} = 1/2$, $k_{21} = 1/2$, $k_{22} = 0$, $a_1 = 1$, $a_2 = 1$, $b_1 = 1/2$, $b_2 = 3/2$, $f_1 = 1$, $f_2 = -5/2$, $g_1 = 2$, $g_2 = 7/2$. Then, $M(k) = N(k) = 1/4$ while $N = 9/2 < 5 = M$.

5. KANTOROVICH'S THEOREM FOR PROBLEM (I)

The finite version of Kantorovich's Theorem [4] can be written as follows:

A feasible solution x_{ij} of Problem (I) is an optimal solution if and only if there exist numbers u_i such that

(5.1) $\quad |u_i - u_j| \leqq c_{ij}$ for each i, j,

(5.2) $\quad u_i - u_j = c_{ij}$ if $x_{ij} > 0$.

We show that this theorem is not valid as it stands. In fact, let us recall Example 2 and let x_{ij} be the optimal solution obtained there. If there exist numbers u_i which satisfy (5.1) and (5.2), then we must have

$$u_1 - u_2 = c_{12} = 20,$$
$$u_3 - u_2 = c_{32} = 1,$$
$$u_1 - u_3 = c_{13} = 1.$$

This is impossible.

In order to give another proof of Kantorovich's Theorem, Kretschmer considered Problem (II) and asserted $N = M$ without any assumption. Notice that $N < M$ in Example 2.

Kantorovich's Theorem was amended by Kantorovich and Rubinstein [5; Theorem 3] in the following form:

THEOREM 4: Assume that conditions (C.1), (C.2) and (C.3) hold. Then a feasible solution x_{ij} of Problem (III) is an optimal solution if and only if there exist numbers u_i which satisfy (5.1) and (5.2).

Under conditions (C.1) and (C.2), the dual problems of Problems (II) and (III) coincide and Theorem 4 is an immediate consequence of the well-known duality theorem applied to Problem (II). Thus, condition (C.3) can be omitted in Theorem 4.

Notice that conditions (C.1), (C.2) and (C.3) hold if and only if the cost c_{ij} is a pseudo-metric, i.e., c_{ij} satisfies conditions (C.1) and (C.2) and the following condition

(C.4) $\qquad c_{ij} \leqslant c_{ik} + c_{kj}$ for all i, j, k.

With the aid of Theorems 2 and 4, we have

THEOREM 5: Assume that conditions (C.1), (C.2) and (C.3) hold. Then a feasible solution x_{ij} of Problem (I) is an optimal solution if and only if there exist numbers u_i which satisfy (5.1) and (5.2).

ACKNOWLEDGMENT

We are indebted to a referee for helpful comments, weakening the original assumptions of Theorem 2, in particular.

REFERENCES

[1] Charnes, A. and W.W. Cooper, *Management Models and Industrial Applications of Linear Programming, I* and *II*, (J. Wiley and Sons, New York, N.Y., 1961).

[2] Dantzig, G.B., *Linear Programming and Extensions*, (Princeton University Press, Princeton, 1963).

[3] Gomory, R.E. and T.C. Hu, "An Application of Generalized Linear Programming to Network Flows," Journal of the Society for Industrial and Applied Mathematics, *10*, 260-283 (1962).

[4] Kantorovich, L.V., "On the Translocation of Masses," Management Science, *5*, 1-4 (1958). (English translation of Doklady Akademii Nauk USSR, *37*, 199-201 (1942).

[5] Kantorovich, L.V. and G. Sh. Rubinstein, "On a Space of Completely Additive Functions," Vestnik Leningrad University, *13*, 52-59 (1958) (Russian).

[6] Kretschmer, K.S., "Programmes in Paired Spaces," Canadian Journal of Mathematics, *13*, 221-238 (1961).

[7] Ryan, M.J., "More on the More for Less Paradox in the Distribution Model," in *Extremal Methods and Systems Analysis, An International Symposium on the Occasion of Professor Abraham Charnes' Sixtieth Birthday*, A.V. Fiacco, K.O. Kortanek (Editors), 275-303, Volume 174 of Lecture Notes in Economics and Mathematical Systems, Managing Editors: M. Beckmann and H.P. Künzi, Springer-Verlag, Berlin-Heidelberg-New York, 1980.

A NETWORK FLOW APPROACH FOR CAPACITY EXPANSION PROBLEMS WITH TWO FACILITY TYPES

Hanan Luss

Bell Laboratories
Holmdel, New Jersey

ABSTRACT

A deterministic capacity expansion model for two facility types with a finite number of discrete time periods is described. The model generalizes previous work by allowing for capacity disposals, in addition to capacity expansions and conversions from one facility type to the other. Furthermore, shortages of capacity are allowed and upper bounds on both shortages and idle capacities can be imposed. The demand increments for additional capacity of any type in any time period can be negative. All cost functions are assumed to be piecewise, concave and nondecreasing away from zero. The model is formulated as a shortest path problem for an acyclic network, and an efficient search procedure is developed to determine the costs associated with the links of this network.

INTRODUCTION

In a previous paper [9], we described a deterministic capacity expansion model for two facility types. The model has a finite number of discrete time periods with known demands for each of the two facilities in any period. At the beginning of each period, facility $i (i = 1, 2)$ may be expanded either by new construction or by converting idle capacity of one facility to accommodate the demand for the other facility.

In this paper, we extend our previous work by allowing for the reduction of facility size through capacity disposals. Furthermore, shortages of capacity are allowed and upper bounds on idle capacities and shortages may be imposed. These generalizations allow us to deal with more realistic situations. Capacity disposals are often initiated due to high holding cost of idle capacity when the cumulative demand decreases over some successive periods. Capacity shortages may be attractive when capacity may be temporarily rented or imported from other sources. Also, in some applications it may be economical to permit temporary shortages and pay a penalty for unsatisfied demand, rather than expanding the facilities at that time. Finally, upper bounds on idle capacity and shortages are usually imposed by management.

The costs incurred include those for construction of new capacity, disposal of existing capacity, conversion, holding of idle capacity, and for having capacity shortages. As in [9], conversion implies physical modification so that the converted capacity becomes an integral part of the new facility and is not reconverted automatically at the end of the period. The capacity expansion policy consists of timing and sizing decisions for new constructions, disposals, and conversions so that the total costs are minimized.

The model is useful for communication network applications, such as the cable sizing problems examined in [9]. Suppose the demands for two cable types is known for the next T periods. Furthermore, suppose the more expensive cable can accommodate both demand types, whereas the cheaper cable can be used only to satisfy its associated demand. Since the construction cost functions are often concave, reflecting economies of scale, it can become attractive to use the more expensive cable for future demand for both cables. Thus, careful planning of the expansion policy is needed. A similar application is the planning of capacity expansion associated with communication facilities which serve digital and analog demands. Other areas of applications include production problems for two substitutable products, and inventory problems of a single product produced and consumed in two separate regions; see [9] for more details.

Many capacity expansion models and closely related inventory models have been developed for the single facility problem with a finite number of discrete time periods. The first such model was proposed by Wagner and Whitin [13] who examined a dynamic version of the economic lot size model. Many authors extended this model; for example, Manne and Veinott [11], Zangwill [16] and Love [8]. Zangwill used a network flow approach, and Love generalized the model to piecewise concave cost functions and bounded idle capacities and shortages.

Several models and algorithms for two facility problems have been developed. Manne [10], Erlenkotter [1,2], Kalotay [5], and Fong and Rao [3] examined models in which it is assumed that converted capacity is reconverted automatically, at no cost, at the end of each period. Kalotay [6], Wilson and Kalotay [14], Merhaut [12], and Luss [9] examined models in which converted capacity is not reconverted automatically at the end of each period.

In Section 1 we describe the generalized model. The algorithm in [9] is extended and used to solve the new model with the additional features described before. In Section 2 a shortest path formulation is presented, and in Section 3 some properties of an optimal solution are identified. These properties are used to compute the costs associated with the links of the network constructed for the shortest path problem. In Section 4 the solution is illustrated by a numerical example, and some final comments are given in Section 5.

1. THE MODEL

The model assumes a finite number of discrete time periods in which the demand increments, new constructions, capacity disposals, and capacity conversions occur instantaneously and simultaneously immediately after the beginning of each period. We define the following notation:

i — index for the two facilities.

t — index for time periods ($t = 1, 2, \ldots, T$) where T is the planning horizon.

r_{it} — the increment of demand for additional capacity of facility i incurred immediately after the beginning of period t. The r_{it}'s may be negative, and for convenience are assumed to be integers.

$$R_i(t_1,t_2) = \sum_{t=t_1}^{t_2} r_{it}, \text{ for } t_1 \leqslant t_2.$$

x_{it} — the amount of new construction ($x_{it} > 0$), or capacity disposal ($x_{it} < 0$), associated with facility i immediately after the beginning of period t.

y_t — the amount of capacity converted immediately after the beginning of period t. $y_t > 0$ ($y_t < 0$) implies that capacity associated with facility 1 (facility 2) is converted to satisfy the demand of the other facility. Once converted, the capacity becomes an integral part of the new facility.

I_{it} — the amount of idle capacity ($I_{it} > 0$), or capacity shortage ($I_{it} < 0$), associated with facility i at the beginning of period t (or equivalently, at the end of period $t-1$, $t = 2, 3, \ldots, T+1$). We assume that initially there is no idle capacity or capacity shortage, that is, $I_{i1} = 0$.

l_{it} — lower bound on I_{it}, that is, the maximum capacity shortage of facility i allowed at the beginning of period t; the l_{it}'s are assumed to be integers and $-\infty \leqslant l_{it} \leqslant 0$.

w_{it} — upper bound on the idle capacity of facility i at the beginning of period t. The w_{it}'s are assumed to be integers and $0 \leqslant w_{it} \leqslant \infty$.

$c_{it}(x_{it})$ — the construction and disposal cost function for facility i at time period t.

$g_t(y_t)$ — the conversion cost function at time period t.

$h_{it}(I_{t,t+1})$ — the cost function associated with idle capacity, or capacity shortage, of facility i carried from period t to period $t+1$.

All cost functions are assumed to be concave from 0 to ∞ and from 0 to $-\infty$, but not necessarily concave over the entire interval $[-\infty, \infty]$. Such functions are called piecewise concave functions, see Zangwill [15]. All cost functions are also assumed to be nondecreasing away from zero; for example, $c_{it}(x_{it})$ is nondecreasing with x_{it} for $x_{it} \geqslant 0$, and nondecreasing with $-x_{it}$ for $x_{it} \leqslant 0$. For convenience, we assume that $c_{it}(0) = g_t(0) = h_{it}(0) = 0$.

The problem can be formulated as follows:

$$(1.1) \qquad \underset{x_{it}, y_t}{\text{Minimize}} \left[\sum_{t=1}^{T} \left(\sum_{i=1}^{2} c_{it}(x_{it}) + h_{it}(I_{t,t+1}) \right) + g_t(y_t) \right]$$

$$(1) \quad \left.\begin{aligned} &(1.2)\ I_{1,t+1} = I_{1t} + x_{1t} - y_t - r_{1t} \\ &(1.3)\ I_{2,t+1} = I_{2t} + x_{2t} + y_t - r_{2t} \\ &(1.4)\ \quad l_{it} \leqslant I_{it} \leqslant w_{it} \\ &(1.5)\ I_{i1} = 0 \\ &(1.6)\ I_{i,T+1} = 0 \end{aligned}\right\} \begin{aligned} &t = 1, 2, \ldots, T \\ &i = 1, 2 \end{aligned}$$

The objective (1.1) is to minimize the total cost incurred over all periods. Equations (1.2) - (1.3) express the idle capacity or capacity shortage $I_{i,t+1}$ as a function of I_{it}, the actions undertaken at period t, x_{it} and y_t, and the demand increments r_{it}. Constraints (1.4) specify the bounds on idle capacities and capacity shortages, and Equation (1.5) is introduced by assumption. Constraint (1.6) $I_{i,T+1} = 0$ implies that idle capacity or capacity shortages are not allowed after period T. Such a constraint is not restrictive since one can add to problem (1) a fictitious

period $T' = T + 1$ with $r_{iT'} = \max_t R_i(1,t) - R_i(1,T)$ (yielding $R_i(1,T') \geq R_i(1,t)\ \forall t$), $I_{iT'} = 0, w_{iT'} = \infty$, and $c_{iT'}(\cdot) = h_{iT'}(\cdot) = g_{T'}(\cdot) = 0$. ($I_{iT'}$ is fixed at zero since no shortages are allowed at the end of period T). This allows us to fix $I_{i,T'+1}$ at zero since then there always exists an optimal solution with $I_{i,T'+1} = 0$. To simplify notation, we assume that period T in formulation (1) is the added fictitious period.

The constraints (1.2) - (1.6) form a nonempty convex set. Since each term of the objective function is nondecreasing away from zero with a finite value at zero, there exists a finite optimal solution. Furthermore, suppose each of the variables x_{it}, y_t, and I_{it} is replaced in formulation (1) by the difference of two nonnegative variables, for example, $x_{it} = x'_{it} - x''_{it}$, where $x'_{it} \geq 0$ represents constructions and $x''_{it} \geq 0$ stands for disposals. In that case, the objective function becomes concave on the entire feasible region; hence, there exists an extreme point optimal solution. From Pages 124-127 in Hu [4], the constraints (1.2) - (1.3) are totally unimodular. Thus, since r_{it}, I_{it} and w_{it} are assumed to be integers, such an extreme point solution consists of integers. In the next sections we describe an algorithm which finds an optimal extreme point solution.

2. A SHORTEST PATH FORMULATION

Since all cost functions are nondecreasing away from zero, it can be shown that there exists an optimal solution in which

$$(2) \qquad |I_{it}| \leq \max_{\tau_1,\tau_2} [R_1(\tau_1,T) + R_2(\tau_2,T)] \equiv b\ \forall i,\ t.$$

However, usually, better bounds than those given by (2) can be assigned. To simplify the presentation, we assume that the lower and upper bounds on the I_{it} variables satisfy $w_{it} \leq b$ and $I_{it} \geq -b$ for all values of i and t.

Generalizing the concept of capacity point in [9], we define a *capacity point* as a period t in which $I_{it} = 0$, or I_{it}, or w_{it} for at least one value of i. Since an extreme point optimal solution consists of integers, the set of capacity points is defined as follows:

$$(3) \quad \begin{cases} (3.1)\ I_{11} = I_{21} = 0 \\ (3.2)\ I_{1t} = I_{1t}, 0, w_{1t} \text{ and } I_{2t} = I_{2t}, 0, w_{2t} \\ (3.3)\ I_{1t} = I_{1t}, 0, w_{1t} \text{ and } I_{2t} = I_{2t} + 1, \ldots, -1, 1, \ldots, w_{2,t-1} \\ (3.4)\ I_{2t} = I_{2t}, 0, w_{2t} \text{ and } I_{1t} = I_{1t} + 1, \ldots, -1, 1, \ldots, w_{1,t-1} \end{cases} \quad t = 2, 3, \ldots, T$$

$$(3.5)\ I_{1,T+1} = I_{2,T+1} = 0.$$

The capacity point values can be conveniently specified by a single parameter α_t. For example, $\alpha_t = 1, 2, \ldots, 9$ can be used to specify the combinations given by (3.2), etc. A complete example of a special case can be found in [9].

The set of capacity points can be limited to those satisfying

$$(4) \quad (4.1)\ I_{1t} + I_{2t} \leq R_1(t,T) + R_2(t,T)$$

$$(4.2)\ I_{1t} + I_{2t} \geq \max_{\tau_1,\tau_2 \leq t-1} [R_1(\tau_1,t-1) + R_2(\tau_2,t-1)].$$

Equation (4.1) states that the total idle capacity at the beginning of period t does not exceed the cumulative demand from period t to T. Equation (4.2) restricts the maximum capacity shortages to the maximum demand increase from any period prior to $t-1$ up to period $t-1$. Clearly, there exists an optimal solution which satisfies (4).

We now describe a shortest path formulation which can be used to solve Problem (1). Let

$d_{uv}(\alpha_u, \alpha_{v+1})$ — the minimal cost during periods $u, u+1, \ldots, v$ associated with an extreme point solution of (1) when u and $v+1$ are two successive capacity points with values defined by α_u and α_{v+1}. More specifically:

$$(5)\quad d_{uv}(\alpha_u, \alpha_{v+1}) = \underset{x_{it}, y_t}{\text{minimum}} \left[\sum_{t=u}^{v} \left(\sum_{i=1}^{2} c_{it}(x_{it}) + h_{it}(I_{i,t+1}) \right) + g_t(y_t) \right]$$

such that

(i) Constraints (1.2) and (1.3) are satisfied for $t = u, u+1, \ldots, v$,

(ii) $l_{it} < I_{it} < w_{it}$ and $I_{it} \neq 0$ for $i = 1, 2$ and $t = u+1, u+2, \ldots, v$,

(iii) I_{1u} and I_{2u} are defined by α_u, and $I_{1,v+1}$ and $I_{2,v+1}$ are defined by α_{v+1},

(iv) x_{it} and y_t for $t = u, u+1, \ldots, v$ satisfy the necessary conditions (to be developed later) for an extreme point solution of (1).

Suppose that all *subproblem values* $d_{uv}(\alpha_u, \alpha_{v+1})$ are known. The optimal solution can then be found by searching for the optimal sequence of capacity points and their associated values. As shown in Figure 1, Problem 1 can be formulated as a shortest path problem for an acyclic network in which the nodes represent all possible values of capacity points. Each node is described by two values (t, α_t) where t is the time period and α_t is the associated capacity point value. From each node (u, α_u) there emanates a directed link to any node $(v+1, \alpha_{v+1})$ for $v \geqslant u$ with an associated cost of $d_{uv}(\alpha_u, \alpha_{v+1})$.

Let C_t be the number of capacity point values at period t. Clearly, $C_1 = C_{T+1} = 1$, and C_t for all other periods can be obtained from Equations (3) and (4). The total number of links N in the shortest path problem is

$$(6)\quad N = \sum_{t=1}^{T} C_t \left(\sum_{j=t+1}^{T+1} C_j \right).$$

Since most of the computational effort is spent on computing the $d_{uv}(\alpha_u, \alpha_{v+1})$ values, it is important to reduce N, if possible. One way, of course, is to reduce the values of C_t through the imposition of appropriate bounds l_{it} and w_{it}.

The shortest path problem can be solved using various algorithms. Since the network is acyclic a simple dynamic programming formulation can be used. Let α_t be described by the set of integers $1, 2, \ldots, C_t$, where $\alpha_t = 1$ represents $I_{1t} = I_{2t} = 0$. Furthermore, let $f_t(\alpha_t)$ be the cost of an optimal policy over periods $t, t+1, \ldots, T$, given that period t is a capacity point, and that I_{1t} and I_{2t} are specified by α_t. The following dynamic programming formulation is then obtained:

$$f_{T+1}(\alpha_{T+1}) = 0, \ \alpha_{T+1} = 1$$

$$(7)\quad f_u(\alpha_u) = \min_{\substack{u \leqslant v \leqslant T \\ 1 \leqslant \alpha_{v+1} \leqslant C_{v+1}}} [d_{uv}(\alpha_u, \alpha_{v+1}) + f_{v+1}(\alpha_{v+1})],$$

$$u = T, T-1, \ldots, 1$$

$$\alpha_u = 1, 2, \ldots, C_u.$$

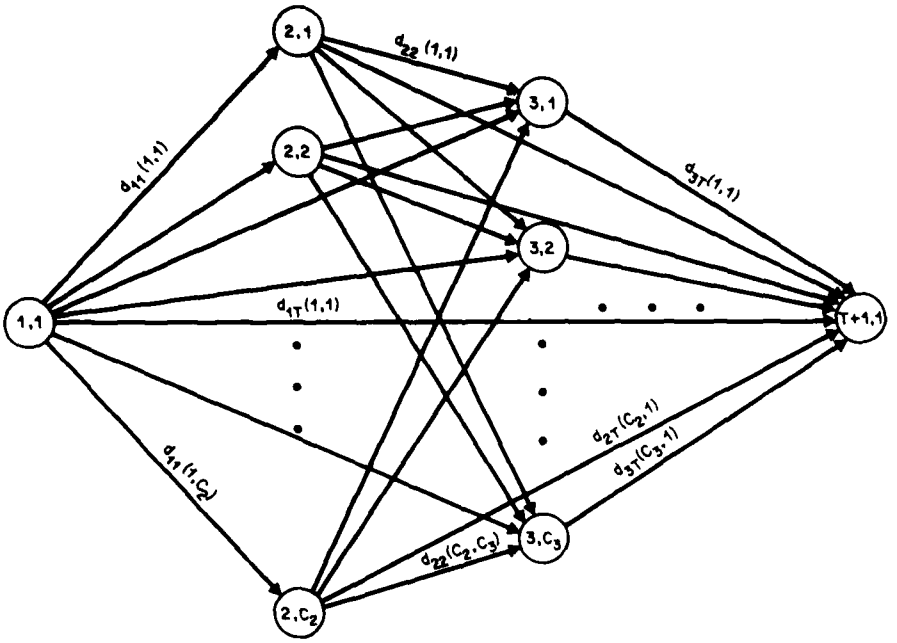


FIGURE 1. The shortest path formulation

The first term of the minimand is the minimum cost of the optimal policy during periods u, $u + 1, \ldots,$ v, given that u and v + 1 are two successive capacity points with values α_u and α_{v+1}. The second term is the optimal cost for periods v + 1, v + 2, ..., T, given α_{v+1}.

3. SOLUTION OF THE SUBPROBLEMS $d_{uv}(\alpha_u, \alpha_{v+1})$

Most of the computational effort is spent on computing the subproblem values. As shown in [9], when $r_{it} \geq 0$, $x_{it} \geq 0$, $l_{it} = 0$ and $w_{it} = \infty$ for all i and t, the subproblems are solved in a trivial manner, however, when the r_{it}'s are allowed to be negative the effort required to solve the subproblems increases significantly. The additional modifications needed to solve the subproblems $d_{uv}(\alpha_u, \alpha_{v+1})$, as defined by (5) for the generalized model, require a more careful analysis than needed in [9], however, the resulting computational effort appears to be about the same.

To compute the subproblem values $d_{uv}(\alpha_u, \alpha_{v+1})$, it is convenient to describe Problem (1) as a single commodity network problem. The network, shown in Figure 2, includes a single source (node 0) with a supply of $R_1(1,T) + R_2(1,T)$. There are 2T additional nodes, each denoted by (i,t) where i specifies the facility and t specifies the time period. At each node (i,t) there is an external demand increment r_{it}, possibly negative. The nodes are connected by links, where the flows along these links represent the constructions, disposals, conversions, idle capacities, and capacity shortages. The flows on each link can be in either direction, and the link direction in Figure 2 indicates positive flows. The nodes are connected by the following links:

- A link from node 0 to each node (i,t) with flow x_{it}. x_{it} is positive if the flow is from node 0 to node (i,t), and negative otherwise.

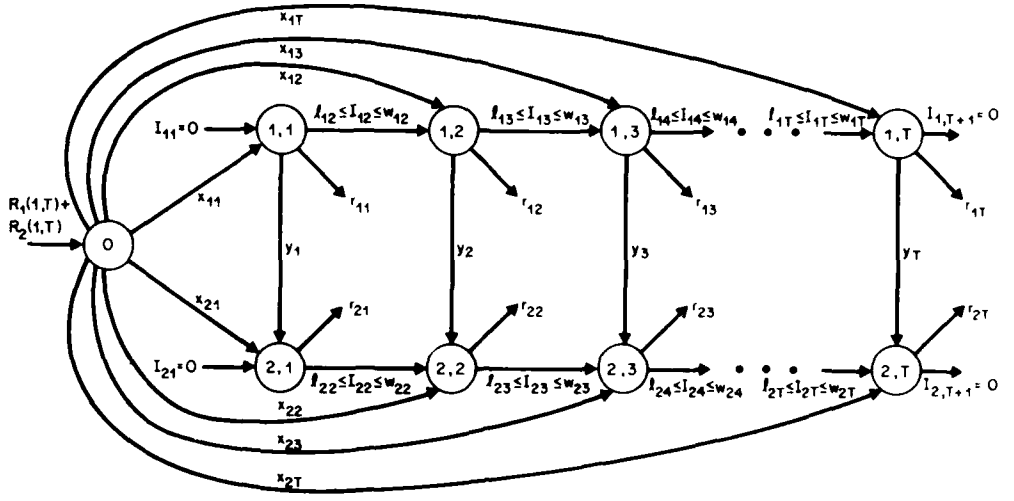


FIGURE 2 A network flow representation of the capacity expansion problem

- A link from each node (i,t) to node $(i,t+1)$ with flow $I_{i,t+1}$. $I_{i,t+1}$ is positive if the flow is from (i,t) to $(i,t+1)$ and negative otherwise.

- A link from each node $(1,t)$ to node $(2,t)$ with flow y_t. y_t is positive if the flow is from node $(1,t)$ to $(2,t)$, and negative otherwise.

As discussed before, we are interested in finding an optimal extreme point solution to a modified version of Problem (1), in which each of the variables x_{it}, y_t, I_{it} is replaced by the difference of two nonnegative variables. It can be shown that a feasible flow in the network given in Figure 2 corresponds to an extreme point solution of Problem (1) modified as described above, if and only if it does not contain any loop with nonzero flows in which all I_{it} flows satisfy $l_{it} < I_{it} < w_{it}$ and $I_{it} \neq 0$.

Concentrating upon a single subproblem, as shown in Figure 3, one may observe that a feasible flow does not contain such loops if and only if the following properties are satisfied:

(8) (8.1) $x_{it_1} x_{it_2} = 0$ $(t_1 \neq t_2)$, $\quad i = 1, 2$

(8.2) $y_{t_1} y_{t_2} = 0$ $(t_1 \neq t_2)$, $\quad u \leqslant t_1, t_2, t_3 \leqslant v$

(8.3) $x_{1t_1} x_{2t_2} y_{t_3} = 0$.

For example, suppose (8.3) is violated and $t_1 \leqslant t_2 \leqslant t_3$, then x_{1t_1}, I_{1,t_1+1}, $\ldots$, I_{1t_3}, y_{t_3}, I_{2t_3}, I_{2,t_3-1}, $\ldots$, I_{2,t_2+1}, x_{2t} form a loop with nonzero flows and all relevant I_{it} values satisfy $l_{it} < I_{it} < w_{it}$ and $I_{it} \neq 0$.

Equation (8) implies that in the optimal solution of $d_{uv}(\alpha_u, \alpha_{v+1})$ there is at most one new construction or disposal for each facility (8.1), and at most one conversion (8.2). Furthermore, if two constructions or disposals (one per facility) are being considered, conversion is then not allowed (8.3).

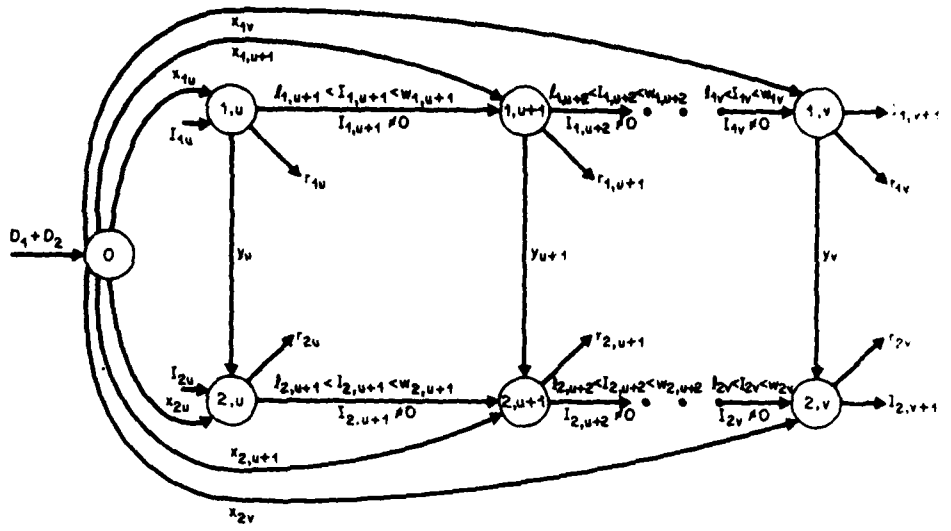


FIGURE 3. A network flow representation of a subproblem

Let D_i be the capacity change of facility i during periods $u, u+1, \ldots, v$, that is:

$$D_i = I_{i,v+1} + R_i(u,v) - I_{iu}, \quad i = 1,2 \tag{9}$$

or, equivalently

$$D_1 = \sum_{t=u}^{v} x_{1t} - y_t \tag{10}$$

$$D_2 = \sum_{t=u}^{v} x_{2t} + y_t.$$

Let t_1 and t_2 be two time periods $u \leqslant (t_1,t_2) \leqslant v$. From the optimal properties (8) shown above, the possible policies associated with an optimal solution to any subproblem $d_{uv}(\alpha_u, \alpha_{v+1})$ can be restricted to three different policies. These policies are summarized in Table 1 below.

To illustrate the table, let us concentrate on the column $D_1 \leqslant 0$ and $D_2 \geqslant 0$. Policy (a) indicates a single disposal of D_1 capacity units of facility 1, and a single construction of D_2 units of facility 2. Policy (b) implies a single construction of $D_1 + D_2$ of facility 1 if $D_1 + D_2 \geqslant 0$, a single disposal of $D_1 + D_2$ of facility 1 if $D_1 + D_2 \leqslant 0$, and a single conversion of D_2 units from facility 1 to facility 2. Obviously, if $D_1 + D_2 = 0$, no constructions or disposals take place, and if $D_2 = 0$, no capacity is converted. Finally, policy (c) consists of a single construction of $D_1 + D_2$ capacity units of facility 2 if $D_1 + D_2 \geqslant 0$, a single disposal of $D_1 + D_2$ units of facility 2 if $D_1 + D_2 \leqslant 0$, and a single conversion of $-D_1$ from facility 1 to facility 2.

The optimal solution of a subproblem $d_{uv}(\alpha_u, \alpha_{v+1})$ is therefore obtained by the following procedure:

(1) For each of the policies (a), (b), and (c) in Table 1, find the optimal values of t_1 and t_2, which minimize $d_{uv}(\alpha_u, \alpha_{v+1})$ as given by Equation (5), while satisfying conditions (i) - (iv) given below Equation (5). If no feasible values of t_1 and t_2 exist, set the value of the corresponding policy to ∞.

TABLE 1. *Possible Policies for Optimal Subproblem Solutions*

Policy \ D_1, D_2		$D_1 \geq 0$ $D_2 \geq 0$	$D_1 \leq 0$ $D_2 \geq 0$	$D_1 \geq 0$ $D_2 \leq 0$	$D_1 \leq 0$ $D_2 \leq 0$
(a)	$x_{1t_1} = D_1,\ x_{1t} = 0\ t \neq t_1$ $x_{2t_2} = D_2,\ x_{2t} = 0\ t \neq t_2$ $y_t = 0\ \forall t$	construction construction	disposal construction	construction disposal	disposal disposal
(b)	$x_{1t_1} = D_1 + D_2,\ x_{1t} = 0\ t \neq t_1$ $y_{t_2} = D_2,\ y_t = 0\ t \neq t_2$ $x_{2t} = 0\ \forall t$	construction conversion from 1 to 2	construction or disposal conversion from 1 to 2	construction or disposal conversion from 2 to 1	disposal conversion from 2 to 1
(c)	$x_{2t_1} = D_1 + D_2,\ x_{2t} = 0\ t \neq t_1$ $y_{t2} = -D_1,\ y_t = 0\ t \neq t_2$ $x_{1t} = 0\ \forall t$	construction conversion from 2 to 1	construction or disposal conversion from 1 to 2	construction or disposal conversion from 2 to 1	disposal conversion from 1 to 2

(2) Choose as the optimal policy the best of those found in Step (1). If none of the policies is feasible, $d_{uv}(\alpha_u, \alpha_{v+1}) = \infty$.

The procedure above may involve spending a significant amount of computation on finding all feasible policies and comparing the costs associated with these policies.

4. A NUMERICAL EXAMPLE

As an illustration, we solve the capacity expansion problem shown in Figure 4 $I_{14} = I_{24} = 0$ by assumption, thus, a fictitious period is not added. The cost functions are given in Table 2 below.

The shortest path formulation is shown in Figure 5. The capacity point values are given inside the nodes in terms of I_{1t} and I_{2t} rather than α_t. Using Equation (4.1), several capacity point values are omitted in periods 2 and 3. Furthermore, all links from period $t = 1$ to periods $t = 3$ and 4 are omitted since there is no feasible solution to the associated subproblems with $I_{12} < I_{12} < w_{12}$ and $I_{12} \neq 0$. The number associated with each link is the optimal solution of the corresponding subproblem. The shortest path is marked by stars.

Consider the subproblem $d_{11}(\alpha_1, \alpha_2)$ where α_1 represents the capacity point value $I_{11} = I_{21} = 0$, and α_2 represents $I_{21} = I_{22} = 0$. By Equation (9), $D_1 = 1$ and $D_2 = 1$. Using the results of Table 1, policy (a) yields $x_{11} = x_{21} = 1$ with a total cost of 68, policy (b) yields $x_{11} = 2$ and $y_1 = 1$ with a cost of 46, and policy (c) yields $x_{21} = 2$ and $y_1 = -1$ with a cost of 45. Hence policy (c) is the optimal one.

To illustrate further, consider $d_{21}(\alpha_2, \alpha_4)$, where α_2 stands for $I_{12} = -1$ and $I_{22} = 0$, and α_4 stands for $I_{14} = I_{24} = 0$, so that $D_1 = 1$ and $D_2 = 0$. From Table 1, policy (a) implies that either $x_{12} = 1$ or $x_{13} = 1$. However, if $x_{13} = 1$ (and $x_{12} = 0$) then $I_{13} = 0$ so that $d_{21}(\cdot) = \infty$. Hence, policy (a) implies $x_{12} = 1$ with construction and holding cost of 43.2. Policy (b) yields the same solution as policy (a), and policy (c) results in $x_{22} = 1$ and $y_2 = -1$ with a total cost of 40.5; hence, policy (c) is optimal for that subproblem.

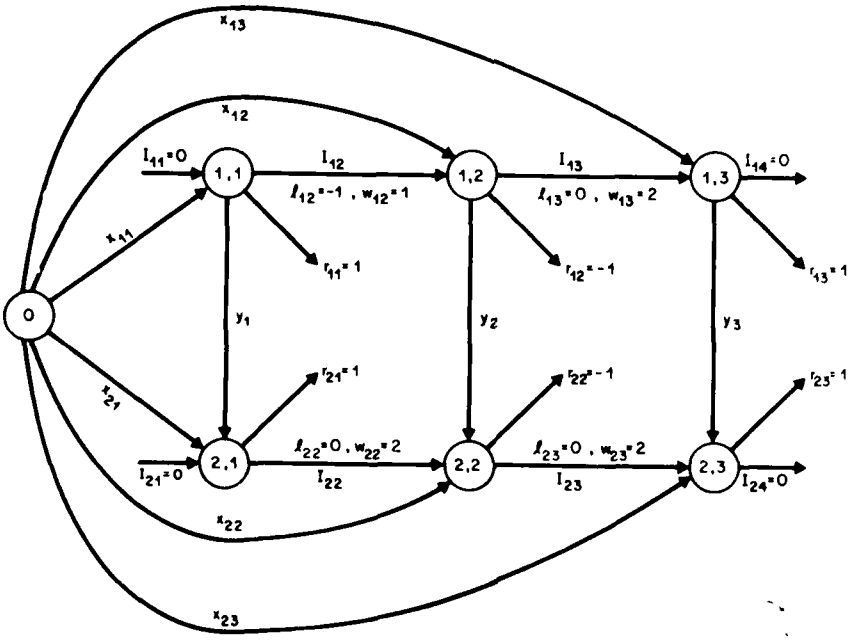


FIGURE 4. A network flow representation of the example

TABLE 2 — *The Cost Functions*

function / argument	$c_{1t}(x_{1t})$	$c_{2t}(x_{2t})$	$h_{it}(I_{i,t+1})$ $i = 1, 2$	$g_t(y_t)$
positive	$(30 + 8x_{1t})0.9^{t-1}$	$(20 + 10x_{2t})0.9^{t-1}$	$5I_{i,t+1}0.9^{t-1}$	0
zero	0	0	0	0
negative	$6 \cdot 0.9^{t-1}$	$5 \cdot 0.9^{t-1}$	$-10I_{i,t+1}0.9^{t-1}$	$-5y_t0.9^{t-1}$

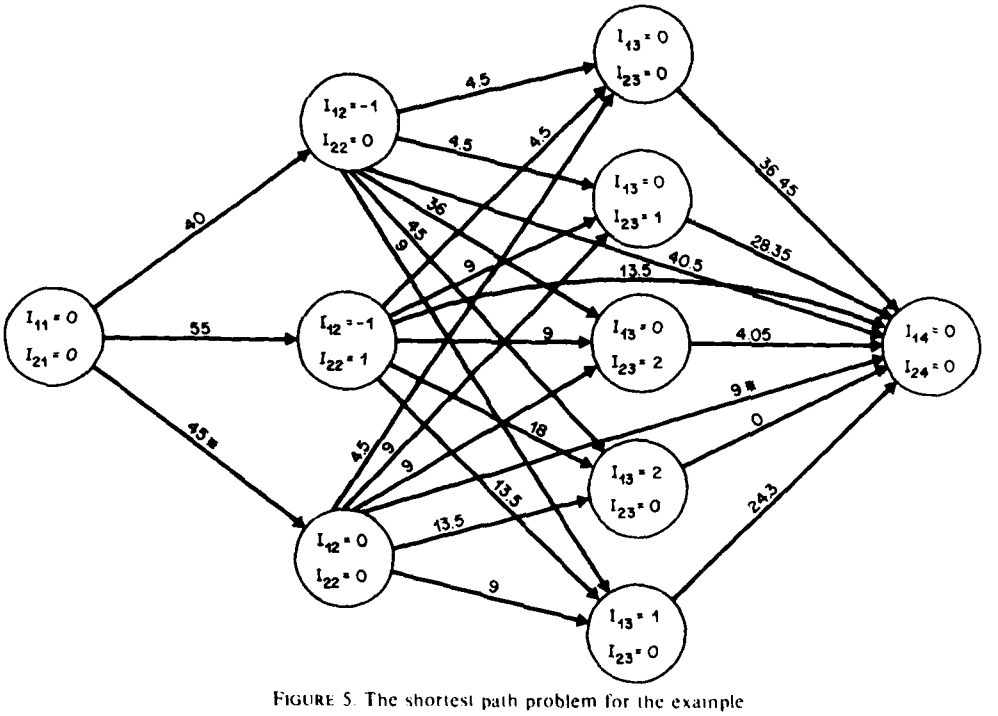


FIGURE 5. The shortest path problem for the example

Finally, consider $d_{23}(\alpha_2, \alpha_4)$ with α_2 standing for $I_{12} = I_{22} = 0$, and α_4 standing for $I_{14} = I_{24} = 0$. From Table 1, since $D_1 = D_2 = 0$, all decision variables are zero in all three policies and the total costs incurred are equal to 9.

After solving the subproblems for all the links of Figure 5, the shortest path can be found using the dynamic programming formulation (7) or any other shortest path algorithm. The shortest path in this example is 54 and consists of two links. The first link connects node $I_{11} = I_{21} = 0$ to node $I_{12} = I_{22} = 0$, and the second link connects node $I_{12} = I_{22} = 0$ to node $I_{14} = I_{24} = 0$. The optimal policy of the entire problem is $x_{21} = 2$, $y_1 = -1$, with all other decision variables x_{it} and y_t being equal to zero.

5. FINAL COMMENTS

This paper generalizes our previous work [9] by allowing for capacity disposals and capacity shortages. Furthermore, bounds on idle capacities and capacity shortages can be imposed. The model is formulated as a shortest path problem in which most of the computational effort is spent on computing the link costs. Using a network flow approach, properties of extreme point solutions are identified. These properties are used to develop an efficient search for the link costs.

Further generalizations may include bounds on new constructions and capacity disposals, and operating costs which depend on the facility type and time period. As shown by several authors, for example Lambrecht and Vander Eecken [7], bounded constructions or disposals complicate considerably even the single facility problem. Introducing operating costs may require major changes in the algorithm since the amount of each capacity type used to satisfy the demand in each period affects the total cost.

Finally, negative costs for disposals (credit for salvage value) can be incorporated for certain cost functions $c_{it}(x_{it})$ for which the optimal solution would be finite. For example, cost functions in which the credit per unit of disposed capacity is always smaller than the construction cost per unit of capacity. In general, however, cost functions $c_{it}(x_{it})$ that are negative for $x_{it} < 0$ may result in an unbounded solution.

REFERENCES

[1] Erlenkotter, D., "Two Producing Areas—Dynamic Programming Solutions," *Investments for Capacity Expansion: Size, Location, and Time Phasing*, 210-227, A. S. Manne, Editor, (MIT Press, Cambridge, Massachusetts, 1967).

[2] Erlenkotter, D., "A Dynamic Programming Approach to Capacity Expansion with Specialization," Management Science, *21*, 360-362 (1974).

[3] Fong, C.O., and M.R. Rao, "Capacity Expansion with Two Producing Regions and Concave Costs," Management Science, *22*, 331-339 (1975).

[4] Hu, T.C., *Integer Programming and Network Flows*, 124-127, (Addison Wesley, Reading, Massachusetts, 1969).

[5] Kalotay, A.J., "Capacity Expansion and Specialization," Management Science, *20*, 56-64 (1973).

[6] Kalotay, A.J., "Joint Capacity Expansion without Rearrangement," Operational Research Quarterly, *26*, 649-658 (1975).

[7] Lambrecht, M. and J. Vander Eecken, "Capacity Constrained Single Facility Lot Size Problem," European Journal of Operational Research, *2*, 132-136 (1978).

[8] Love, S.F., "Bounded Production and Inventory Models with Piecewise Concave Costs," Management Science, *20*, 313-318 (1973).

[9] Luss, H., "A Capacity-Expansion Model for Two Facilities," Naval Research Logistics Quarterly, *26*, 291-303 (1979).

[10] Manne, A.S., "Two Producing Areas—Constant Cycle Time Policies," *Investments for Capacity Expansion: Size, Location, and Time Phasing*, 193-209, A.S. Manne, Editor, (MIT Press, Cambridge, Massachusetts, 1967).

[11] Manne, A.S. and A.F. Veinott, Jr., "Optimal Plant Size with Arbitrary Increasing Time Paths of Demand," *Investments for Capacity Expansion: Size, Location, and Time Phasing*, 178-190, A.S. Manne, Editor, (MIT Press. Cambridge Massachusetts, 1967).

[12] Merhaut, J.M., "A Dynamic Programming Approach to Joint Capacity Expansion without Rearrangement," M. Sc. Thesis, Graduate School of Management, University of California, Los Angeles, California (1975).

[13] Wagner, H.M. and T.M. Whitin, "Dynamic Version of the Economic Lot Size Model," Management Science, *5*, 89-96 (1958).

[14] Wilson, L.O., and A.J. Kalotay, "Alternating Policies for Nonrearrangeable Networks," INFOR, *14*, 193-211 (1976).

[15] Zangwill, W.I., "The Piecewise Concave Function," Management Science, *13*, 900-912 (1967).

[16] Zangwill, W.I., "A Backlogging Model and a Multiechelon Model for a Dynamic Economic Lot Size Production System—A Network Approach," Management Science, *15*, 506-527 (1969).

SOLVING MULTIFACILITY LOCATION PROBLEMS INVOLVING EUCLIDEAN DISTANCES*

Paul Calamai

Department of Systems Design
University of Waterloo
Waterloo, Ontario, Canada

Christakis Charalambous

Department of Electrical Engineering
Concordia University
Montreal, Quebec, Canada

ABSTRACT

This paper considers the problem of locating multiple new facilities in order to minimize a total cost function consisting of the sum of weighted Euclidean distances among the new facilities and between the new and existing facilities, the locations of which are known. A new procedure is derived from a set of results pertaining to necessary conditions for a minimum of the objective function. The results from a number of sample problems which have been executed on a programmed version of this algorithm are used to illustrate the effectiveness of the new technique.

1. BACKGROUND

It was as early as the 17th century that mathematicians, notably Fermat, were concerned with what are now known as single facility location problems. However, it was not until the 20th century that normative approaches to solving symbolic models of these and related problems were addressed in the literature. Each of these solution techniques concerned themselves with determining the location of a new facility, or new facilities, with respect to the location of existing facilities so as to minimize a cost function based on a weighted interfacility distance measure.

If one studies a list of references to the work done in the past decade involving facility location problems it becomes readily apparent that there exists a strong interdisciplinary interest in this area within the fields of operations research, management science, logistics, economics, urban planning and engineering. As a result, the term "facility" has taken on a very broad connotation in order to suit applications in each of these areas. Francis and Goldstein [4] provide a fairly recent bibliography of the facility location literature. One of the most complete classifications of these problems is provided in a book by Francis and White [5].

*This work was supported by the National Research Council of Canada under Grant A4414 and by an Ontario Graduate Scholarship awarded to Paul Calamai.

This paper concerns itself with the development of an algorithm for solving one particular problem in the area of facility location research. The problem involves multiple new facilities whose locations, the decision variables, are points in E_2 space. The quantitative objective is to minimize the total cost function consisting of the sum of weighted Euclidean distances among new facilities and between new and existing facilities. The weights are the constants of proportionality relating the distance travelled to the costs incurred. It is assumed that the problem is "well structured" [3].

The Euclidean distance problem for the case of single new facilities was addressed by Weiszfeld [13], Miehle [10], Kuhn and Kuenne [8], and Cooper [1] to name a few. However, it was not until the work of Kuhn [7] that the problem was considered completely solved. A computational procedure for minimizing the Euclidean multifacility problem was presented by Vergin and Rogers [12] in 1967; however, their techniques sometimes give suboptimum solutions. Two years later, Love [9] gave a scheme for solving this problem which makes use of convex programming and penalty function techniques. One advantage to this approach is that it considers the existence of various types of spatial constraints. In 1973 Eyster, White and Wierwille [2] presented the hyperboloid approximation procedure (HAP) for both rectilinear and Euclidean distance measures which extended the technique employed in solving the single facility problem to the multifacility case. This paper presents a new technique for solving continuous unconstrained multifacility location problems involving Euclidean distances.

2. PROBLEM FORMULATION

The continuous unconstrained multifacility location problem involving the l_p distance measure can be stated as follows:

Find the point $X^{*T} = (X^{*T}_1, \ldots, X^{*T}_n)$ in E_{2n} to

$$\text{(P1)} \qquad \text{minimize } f(X) = \sum_{1 \leqslant j < k \leqslant n} v_{jk} \, ||X_j - X_k||_p + \sum_{j=1}^{n} \sum_{i=1}^{m} w_{ji} \, ||X_j - A_i||_p$$

where

$n \triangleq$ number of new facilities (NF's).

$m \triangleq$ number of existing facilities (*EF*'s).

$X_j^T = (x_{j1} \; x_{j2}) \triangleq$ vector location of NF_j in E_2, $j = 1, \ldots, n$.

$A_i^T = (a_{i1} \; a_{i2}) \triangleq$ vector location of EF_i in E_2, $i = 1, \ldots, m$.

$v_{jk} \triangleq$ nonnegative constant of proportionality relating the l_p distance between NF_j and NF_k to the cost incurred $1 \leqslant j < k \leqslant n$.

$w_{ji} \triangleq$ nonnegative constant of proportionality relating the l_p distance between NF_j and EF_i to the cost incurred $1 \leqslant j \leqslant n$, $1 \leqslant i \leqslant m$.

$||X_j - X_k||_p = \{|x_{j1} - x_{k1}|^p + |x_{j2} - x_{k2}|^p\}^{1/p} \triangleq l_p$ distance between NF_j and NF_k.

$||X_j - A_i||_p = \{|x_{j1} - a_{i1}|^p + |x_{j2} - a_{i2}|^p\}^{1/p} \triangleq l_p$ distance between NF_j and EF_i.

Note that we make the assumption that $v_{jk} = v_{kj}$ for $j,k = 1, \ldots, n$. Substituting $p = 1$ and $p = 2$ in Problem P1 respectively yields the rectilinear distance problem and the Euclidean distance problem.

For the purpose of this paper Euclidean distance will be the measure used between facilities located as points in E_2 space. The objective function becomes

$$\underset{X}{\text{minimize}}\ f(X) = \sum_{1 \leq j < k \leq n} v_{jk} \{(x_{j1} - x_{k1})^2 + (x_{j2} - x_{k2})^2\}^{1/2}$$

$$+ \sum_{j=1}^{n} \sum_{i=1}^{m} w_{ji} \{(x_{j1} - a_{i1})^2 + (x_{j2} - a_{i2})^2\}^{1/2}. \tag{P2}$$

The techniques presented in this paper can also be used for problems involving facilities located in three-dimensional space.

3. NEW FACILITY CATEGORIZATION

If we consider a current solution to Problem P2 we can think of each new facility as being in one of the following distinct categories:

(1) Unique Point (*UP*)

A new facility in this category occupies a location that differs from all other facility locations.

(2) Coinciding Point (*CP*)

A new facility in this category occupies a location that coincides with the location of an existing facility but differs from the current locations of all other new facilities. Thus, each new facility in this category has associated with it some existing facility which has the same vector location.

(3) Unique Clusters ($UC_1, \ldots, UC_{NUC}$)

All new facilities in the kth unique cluster ($k = 1, \ldots, NUC$) occupy the same vector location. This location is distinct from all existing facility locations as well as the current locations of new facilities that are not classified in this cluster.

(4) Coinciding Clusters ($CC_1, \ldots, CC_{NCC}$)

All new facilities categorized in the kth coinciding cluster ($k = 1, \ldots, NCC$) occupy the same vector location. This location coincides with the location of some existing facility and differs from the current locations of all new facilities that are not classified in this cluster. Each of these coinciding clusters of new facilities is therefore associated with some existing facility with which it shares a location.

If we define the index sets $J \triangleq \{1, \ldots, n\}$ and $I \triangleq \{1, \ldots, m\}$ and let the subsets $UC_0 = CC_0 = \phi$ then the categorization can be restated as follows:

Partition the set J into the subsets UP, CP, $UC_1, \ldots, UC_{NUC}$, $CC_1, \ldots, CC_{NCC}$ where

$$UP = \{\forall_j \in J | A_i \neq X_j \neq X_k; \forall i \in I, \forall k \in J - \{j\}\} \tag{3.1}$$

$$CP = \{\forall_j \in J | A_{i_j} = X_j \neq X_k; i_j \in I, \forall k \in J - \{j\}\} \tag{3.2}$$

for $\alpha = 1, \ldots, NUC$

$$UC_\alpha \triangleq \left\{\forall_j \in J - \bigcup_{l=0}^{\alpha-1} UC_l | A_i \neq X_j = X_k; \forall i \in I, k \in J - \{j\} - \bigcup_{l=0}^{\alpha-1} UC_l\right\} \tag{3.3}$$

for $\beta = 1, \ldots, NCC$

$$CC_\beta \triangleq \left\{ \forall j \in J - \bigcup_{l=0}^{\beta-1} CC_l | A_{i_\beta} = X_j = X_k;\ i_\beta \in I,\ k \in J - \{j\} - \bigcup_{l=0}^{\beta-1} CC_l \right\} \tag{3.4}$$

$NUC \triangleq$ number of unique clusters.

$NCC \triangleq$ number of coinciding clusters.

Note that

(a) New facility j coincides with existing facility i_j for $j \in CP$ (from 3.2).

(b) The new facilities in cluster β coincide with existing facility i_β for $\beta = 1, \ldots, NCC$ (from 3.4).

In order to use this notation for the derivation of the new algorithm given in the next section define a unit vector D in E_{2n} as follows:

$$D^T = \{D_1^T, \ldots, D_n^T\}$$

where

$$D_j^T = [d_{j1}\ d_{j2}], \quad j = 1, \ldots, n \tag{3.5}$$

and

$$||D||_2 = 1.$$

4. THE DIRECTIONAL DERIVATIVE

Using the notation given in the last section we can write the directional derivative of the objective function at X in the direction D in the following useful manner:

$$\begin{aligned} d_D f(X) &= \lim_{\lambda \to 0^+} \frac{f(X + \lambda D) - f(X)}{\lambda} \\ &= \sum_{j \in UP} [G_j \cdot D_j] \\ &\quad + \sum_{j \in CP} [G_j \cdot D_j + w_{ji_j} ||D_j||_2] \\ &\quad + \sum_{\alpha=1}^{NUC} \left\{ \sum_{j \in UC_\alpha} \left[G_j \cdot D_j + \sum_{\substack{k \in UC_\alpha \\ k > j}} v_{jk} ||D_j - D_k||_2 \right] \right\} \\ &\quad + \sum_{\beta=1}^{NCC} \left\{ \sum_{j \in CC_\beta} \left[G_j \cdot D_j + \sum_{\substack{k \in CC_\beta \\ k > j}} v_{jk} ||D_j - D_k||_2 + W_{ji_\beta} ||D_j|| \right] \right\} \end{aligned} \tag{4.1}$$

where

$$G_j = \sum_{k \neq j} \frac{v_{jk}(X_j - X_k)}{||X_j - X_k||_2} + \sum_{i \in I} \frac{w_{ji}(X_j - A_i)}{||X_j - A_i||_2} \quad \forall j \in UP \tag{4.2a}$$

$$G_j = \sum_{k \neq j} \frac{v_{jk}(X_j - X_k)}{||X_j - X_k||_2} + \sum_{\substack{i \in I \\ \neq i_j}} \frac{w_{ji}(X_j - A_i)}{||X_j - A_i||_2} \quad \forall j \in CP \tag{4.2b}$$

$$G_j = \sum_{k \notin UC_\alpha} \frac{v_{jk}(X_j - X_k)}{||X_j - X_k||_2} + \sum_{i \in I} \frac{w_{ji}(X_j - A_i)}{||X_j - A_i||_2} \qquad \forall_j \in UC_\alpha \quad \alpha = 1, \ldots, NUC \tag{4.2c}$$

$$G_j = \sum_{k \notin CC_\beta} \frac{v_{jk}(X_j - X_k)}{||X_j - X_k||_2} + \sum_{\substack{i \in I \\ \neq i_\beta}} \frac{w_{ji}(X_j - A_i)}{||X_j - A_i||_2} \qquad \forall_j \in CC_\beta \quad \beta = 1, \ldots, NCC \tag{4.2d}$$

It should be noted that in each case, the expression for G_j is the gradient of that part of the objective function $f(X)$, which is differentiable with respect to X_j. In the case where $j \in UP$, the expression is the exact gradient with respect to X_j; in all other cases, the expression for G_j can be considered a pseudo-gradient of $f(X)$ with respect to X_j.

Since $f(X)$ is a convex function, the point X^* in E_{2n} is a minimum for this function if and only if the directional derivative $d_D f(X^*)$ is nonnegative for all unit vectors D in E_{2n}. This fact will be used in the next section.

5. NECESSARY CONDITIONS FOR OPTIMALITY

THEOREM 1: If the following conditions are not satisfied at the point X in E_{2n} then the directional derivative given by expression (4.1) will be negative for some unit vector D in E_{2n}.

(5.1) (1) $||G_j||_2 = 0 \qquad \forall_j \in UP$

(5.2) (2) $||G_j||_2 \leqslant w_{ji_j} \qquad \forall_j \in CP$

(5.3) (3) for $\alpha = 1, \ldots, NUC$

$$||\sum_{j \in S} G_j||_2 \leqslant \sum_{j \in S} \sum_{k \in [UC_\alpha - S]} v_{jk} \qquad \forall S \subset UC_\alpha$$

(5.4) (4) for $\beta = 1, \ldots, NCC$

$$||\sum_{j \in T} G_j||_2 \leqslant \sum_{j \in T} \left[\left(\sum_{k \in [CC_\beta - T]} v_{jk} \right) + w_{ji_\beta} \right] \qquad \forall T \subset CC_\beta$$

PROOF: The proofs for conditions 1 and 2 are obvious.

For 3) set $\begin{cases} D_j = R & \text{for } j \in S \\ D_j = 0 & \text{for } j \notin S \end{cases}$

$$\text{then } d_D f(X) = \sum_{j \in S} G_j \cdot R + \sum_{j \in S} \sum_{k \in [UC_\alpha - S]} v_{jk} ||R||_2$$

$$= ||R||_2 \left\{ ||\sum_{j \in S} G_j||_2 \cos\theta + \sum_{j \in S} \sum_{k \in [UC_\alpha - S]} v_{jk} \right\}.$$

Therefore, $d_D f(X) \geqslant 0 \quad \forall D$ only if

$$||\sum_{j \in S} G_j||_2 \leqslant \sum_{j \in S} \sum_{k \in [UC_\alpha - S]} v_{jk} \qquad \forall S \subset UC_\alpha.$$

The proof for condition 4 is similar.

6. UPDATE FORMULAS

As a result of the preceeding optimality conditions the following update formulas are constructed:

CASE 1: If $\exists j \in UP$ such that $||G_j|| \neq 0$ then the direction of steepest-ascent in the subspace defined by X_j is $\hat{G}_j = G_j$. We therefore use the following update formula for X_j:

$$X_j \leftarrow X_j - \lambda_j \hat{G}_j$$

where

$$\lambda_j = \left[\sum_{k \neq j} \frac{v_{jk}}{||X_j - X_k||_2} + \sum_{i \in I} \frac{w_{ji}}{||X_j - A_i||_2} \right]^{-1}. \tag{6.1}$$

CASE 2: If $\exists j \in CP$ such that $||G_j||_2 > w_{ji_j}$ then the direction of steepest-ascent in the subspace defined by X_j is $\hat{G}_j = G_j$. We therefore use the following update formula for X_j:

$$X_j \leftarrow X_j - \lambda_j \hat{G}_j$$

where

$$\lambda_j = \left[\sum_{k \neq j} \frac{v_{jk}}{||X_j - X_k||_2} + \sum_{\substack{i \in I \\ \neq i_j}} \frac{w_{ji}}{||X_j - A_i||_2} \right]^{-1}. \tag{6.2}$$

CASE 3: If $\exists S \subset UC_\alpha, \alpha = 1, \ldots, NUC$, such that

$$||\sum_{j \in S} G_j||_2 > \sum_{j \in S} \sum_{k \in [UC_\alpha - S]} v_{jk}$$

then the direction of steepest-ascent in the subspace defined by the subset cluster is $\hat{G}_S = \sum_{j \in S} G_j$. We therefore use the following update formula:

$$\forall j \in S \qquad X_j \leftarrow X_j - \lambda_S \hat{G}_S$$

where

$$\lambda_S = \left[\sum_{j \in S} \left[\sum_{k \notin UC_\alpha} \frac{v_{jk}}{||X_j - X_k||} + \sum_{i} \frac{w_{ji}}{||X_j - A_i||} \right] \right]^{-1}. \tag{6.3}$$

CASE 4: If $\exists T \subset CC_\beta, \beta = 1, \ldots, NCC$, such that

$$||\sum_{j \in T} G_j||_2 > \sum_{j \in T} \left[\left[\sum_{k \in [CC_\beta - T]} v_{jk} \right] + w_{ji_\beta} \right]$$

then the direction of steepest-ascent in the subspace defined by the subset cluster is $\hat{G}_T = \sum_{j \in T} G_j$. We therefore use the following update formula:

$$\forall j \in I \qquad X_j \leftarrow X_j - \lambda_T \hat{G}_T$$

where

$$\lambda_T = \left[\sum_{j \in T}\left[\sum_{k \notin CC_\beta} \frac{v_{jk}}{||X_j - X_k||_2} + \sum_{\substack{i \in I \\ \neq i_\beta}} \frac{w_{ji}}{||X_j - A_i||_2}\right]\right]^{-1}. \tag{6.4}$$

In each result, the expression for lambda (λ) can be considered a weighted harmonic mean [8] of the interfacility distance terms appearing in the equation for the gradient (Case 1) or pseudo-gradients (Cases 2 through 4).

7. A NEW ALGORITHM

Using the results derived in the preceeding section the following algorithm can be used to solve Problem P2:

(1) Find a current solution X in E_{2n}.

(2) Try to obtain a better solution by moving single new facilities by using Cases 1 and 2.

(3) For $\alpha = 1, \ldots, NUC$ try to obtain a better solution by applying the special form of Case 3 where $|S| = 1$ (to move single new facilities) or, if this fails, applying the special form of Case 3 where $|S| = |UC_\alpha|$ (to move entire clusters of new facilities). If successful, return to Step 2.

(4) For $\beta = 1, \ldots, NCC$ try to obtain a better solution by applying the special form of Case 4 where $|T| = 1$ (to move single new facilities) or, if this fails, applying the special form of Case 4 where $|T| = |CC_\beta|$ (to move entire clusters of new facilities). If successful, return to Step 2.

(5) Try to obtain a better solution by moving subset clusters using Cases 3 and 4. If an improvement is made, return to Step 2.

8. REMARKS ON IMPLEMENTATION

The following rules were used in implementing the algorithm described in the last section.

(a) New facility j and new facility k were considered "clustered" if:

$$||X_j||_2 + ||X_k||_2 < \epsilon_1 \qquad 1 \leqslant j < k \leqslant n \tag{8.1a}$$

or

$$\frac{2 \cdot ||X_j - X_k||_2}{||X_j||_2 + ||X_k||_2} < \epsilon_1 \qquad 1 \leqslant j < k \leqslant n \tag{8.1b}$$

where $\epsilon_1 \triangleq$ inputted cluster tolerance.

(b) New facility j and existing facility i were considered "coinciding" if:

$$||X_j||_2 + ||A_i||_2 < \epsilon_1 \qquad j = 1, \ldots, n;\ i = 1, \ldots, m \tag{8.2a}$$

or

$$\frac{2 \cdot ||X_j - A_i||_2}{||X_j||_2 + ||A_i||_2} < \epsilon_1 \qquad j = 1, \ldots, n;\ i = 1, \ldots, m \tag{8.2b}$$

where $\epsilon_1 \triangleq$ inputted cluster tolerance,

(c) The update formulas were used only if:

$$\lambda \|\hat{G}\|_2 > \epsilon_2 \tag{8.3}$$

where $\epsilon_2 \triangleq$ inputted step tolerance. This helped avoid the possibility of repeatedly taking small steps. However, the step tolerance is reduced prior to the termination of the algorithm as outlined by the next rule.

(d) In order to ensure optimality, the following check is made prior to executing Step 5 of the algorithm:

$$[f(X^{(h-1)}) - f(X^{(h)})] * 100 \leqslant \epsilon_3 * f(X^{(h-1)}) \tag{8.4}$$

where $\epsilon_3 \triangleq$ inputted function tolerance.

If this condition is not satisfied, the step tolerance (ϵ_2) is reduced and the algorithm restarted at Step 2.

9. DISCUSSION

The new algorithm has the following properties:

(a) It makes full use of the structure of the facility location problem thus avoiding the need for any background in related nonlinear programming areas.

(b) The actual objective function, and not an approximation to it, is minimized at each step in the algorithm.

(c) The stepsize used in this algorithm may not be "optimal" when compared with stepsizes obtained from line-search techniques. However, the use of this stepsize has the following advantages: a) ease of computation, b) maintenance of location problem structure, and c) reduced computation time per update.

(d) Although Step 5 in the algorithm is combinatorial in complexity, very little computational work is necessary. This is a result of the fact that all the information needed for this step has already been computed and stored in previous steps.

(e) The algorithm is similar to the technique devised by Kuhn for solving the single-facility location problems with Euclidean distances [7] and the method devised by Juel and Love [6] for the multifacility location problem with rectilinear distances. This makes the algorithm attractive to those with experience with these methods.

(f) Currently, there is no rigorous proof that this algorithm converges. In 1973, Kuhn [7] completed the proof of convergence for a similar scheme, introduced by Weiszfeld [13] in 1937, for the case of single new facilities. Based on computational experience and on the fact that the algorithm is designed to minimize the objective function in all new facility subspaces, it is likely that the algorithm always converges.

(g) The main disadvantage of the algorithm is that the order in which each of the subspaces is checked is, currently, not optimal. A method, based on projections, that would allow us to determine "a priori" which subspace to update, is now being investigated.

(h) Most existing methods for solving the multifacility problem lack any consideration of the existence of constraints on the solution space [9]. This is also true of the new method outlined in this paper, however, the addition of constraints should not present a problem to the projection technique.

(i) It has yet to be proven that the necessary conditions for optimality for Problem P2 given by Equations (5.1) through (5.4), are also sufficient.

10. COMPUTATIONAL EXPERIENCE

The performance of the algorithm described in this paper (MFLPV1) was tested against the hyperboloid approximation procedure (HAP) described in Eyster, White and Wierwille [2] and a modified hyperboloid approximation procedure (MHAP) suggested by Ostresh [11].

Two parameters were used as a basis of comparison: 1) the number of new facility location updates needed to reach optimality, and 2) the required CPU time in minutes. In the case of program MFLPV1, two counts were considered necessary for specifying the first parameter. The first count represented the number of "attempted" updates (excluding those updates from Step 5 of the algorithm). The second count represented the number of "successful" updates. The reason for excluding the number of attempted updates from Step 5 of the algorithm is simply this: computationally, very little work is done at this step in the procedure.

Six problems were used for the comparison; the first three were taken from [5] (#5.23, #5.7 and #5.6 respectively), the fourth appears in [2] and the last two problems summarized in Tables 1 and 2, are the authors.

HAP and MHAP were both executed using two different initial hyperbolic constants ϵ^{-1} for these problems in order to emphasize the significance of this parameter to the performance of these algorithms. The stopping criteria used in each case was the same as that outlined in the paper introducing HAP [2]. Unless otherwise specified, program MFLPV1 also made use of the following data.

(1) $\epsilon_1 \triangleq$ cluster tolerance = 0.01 (from Equations (8.1) and (8.2)).

(2) $\epsilon_2 \triangleq$ step tolerance = 0.05 (from Equation (8.3))

(3) $\epsilon_3 \triangleq$ function tolerance = 0.01 (from Equation (8.4)).

The results of these tests are summarized in Table 3. The numbers in this table represent the total new facility updates required to reach optimality. The numbers in brackets ( ), under the column headed MFLPV1, represent the number of successful updates whereas the unbracketed numbers in these columns represent the number of attempted updates. The following observations and comments can be made about the results summarized in this table.

(a) In all but Problem 5, the number of attempted updates required to reach optimality using MFLPV1 is less than the number of updates required by HAP and MHAP. These numbers are directly comparable.

(b) The new procedure (MFLPV1) used considerably less CPU time in solving the six problems than did HAP and MHAP.

TABLE 1 — *Input Parameters for Problem 5*

i	a_{i1}	a_{i2}
1	0.0	0.0
2	2.0	4.0
3	6.0	2.0
4	6.0	10.0
5	8.0	8.0

(a) EF Locations

j	$x_{j1}^{(0)}$	$x_{j2}^{(0)}$
1	0.0	0.0
2	0.0	0.0
3	6.0	10.0
4	1.0	3.0
5	6.0	10.0
6	8.0	8.0
7	2.0	4.0
8	2.0	4.0
9	6.0	10.0

(b) Initial NF Locations

j \ i	1	2	3	4	5
1	1.0	1.0	1.0	1.0	1.0
2	1.0	1.0	1.0	1.0	1.0
3	1.0	1.0	1.0	1.0	1.0
4	1.0	1.0	1.0	1.0	1.0
5	1.0	1.0	1.0	1.0	1.0
6	1.0	1.0	1.0	1.0	1.0
7	1.0	1.0	1.0	1.0	1.0
8	1.0	1.0	1.0	1.0	1.0
9	1.0	1.0	1.0	1.0	1.0

(c) w_{ij} Weights

j \ k	1	2	3	4	5	6	7	8	9
1	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	
2		1.0	1.0	1.0	1.0	1.0	1.0	1.0	
3			1.0	1.0	1.0	1.0	1.0	1.0	
4				1.0	1.0	1.0	1.0	1.0	
5					1.0	1.0	1.0	1.0	
6						1.0	1.0	1.0	
7							1.0	1.0	
8								1.0	
9									

(d) v_{jk} Weights

TABLE 2 — *Input Parameters for Problem 6*

i	a_{i1}	a_{i2}
1	2.0	5.0
2	10.0	20.0
3	10.0	10.0

(a) EF Locations

j	$x_{j1}^{(0)}$	$x_{j2}^{(0)}$
1	5.0	15.0
2	5.0	15.0

(b) Initial NF Locations

j \ i	1	2	3
1	0.16	0.56	0.16
2	0.16	0.56	0.16

(c) w_{ij} Weights

j \ k	1	2
1		1.5
2		

(d) v_{jk} Weights

TABLE 3 — *Comparative Test Results for Six Problems*

#	MFLPV1	$\epsilon^{(0)} = 10^{0}$		$\epsilon^{(0)} = 10^{-4}$		X^*	$f(X^*)$
		HAP	MHAP	HAP	MHAP		
1	564 (77)	1661	1381	2027	1407	(1.0,0.0) (1.0,0.0) (1.0,0.0) (2.0,0.0) (2.0,0.0)	38.990
2	148 (34)	647	546	4641	2281	(10.0,20.0) (10.0,20.0)	186.798
3	63 (16)	87	70	770	197	(8.0,7.0) (8.0,7.0)	43.351
4	31 (15)	45	45	45	45	(2.832,2.692 (5.096,6.351)	67.250
5	223 (40)	142	114	1763	975	(4.045,4.281) (4.045,4.281) (4.045,4.281) (4.045,4.281) (4.045,4.281) (4.045,4.281) (4.045,4.281) (4.045,4.281) (4.045,4.281)	201.878
6	63 (7)	242	164	3743	1869	(10.0,20.0) (10.0,20.0)	8.540
TOTAL	1092 (189)	2824	2320	12989	6774		
CPU	0.07	0.45	0.50	1.88	1.48		

(c) Five of the six problems have solutions at cluster points. This appears to be the case in many other problems. This suggests that methods using clustering information, such as MFLPV1, will perform better than methods that disregard this information.

10. CONCLUDING REMARKS

To date, many of the methods designed for solving the multifacility location problem have been either poorly structured, suboptimal or haphazard. In this paper, a new method is developed for solving the multifacility location problem involving Euclidean distances. This new method can easily be extended to accommodate problems involving item movements that are other than Euclidean. Computational experience shows that this method outperforms techniques currently in use. In addition, the proposed method takes full advantage of the structure of the location problem.

Most current techniques used for solving location problems, including those proposed in this paper, are designed to minimize an unconstrained objective function. This is an incomplete treatment since most practical problems involve some form of spatial constraints. It is

proposed that these constraints be handled and the performance of the algorithm improved through the use of projection techniques. This approach is currently being investigated by the authors.

BIBLIOGRAPHY

[1] Cooper, L., "Location-Allocation Problems," Operations Research, *11*, 331-344 (1963).

[2] Eyster, J.W., J.A. White and W.W. Wierwille, "On Solving Multifacility Location Problems Using a Hyperboloid Approximation Procedure," American Institute of Industrial Engineers Transactions, *5*, 1-6 (1973).

[3] Francis, R.L. and A.V. Cabot, "Properties of a Multifacility Location Problem Involving Euclidean Distances," Naval Research Logistics Quarterly, *19*, 335-353 (1972).

[4] Francis, R.L. and J.M. Goldstein, "Location Theory: A Selective Bibliography," Operations Research, *22*, 400-410 (1974).

[5] Francis, R.L. and J.A. White, "*Facility Layout and Location: An Analytic Approach*," Prentice-Hall, Englewood Cliffs, New Jersey (1974).

[6] Juel, H. and R.F. Love, "An Efficient Computational Procedure for Solving the Multi-Facility Rectilinear Facilities Location Problem," Operational Research Quarterly, *27*, 697-703 (1976).

[7] Kuhn, H.W., "A Note on Fermat's Problem," Mathematical Programming, *4*, 98-107 (1973).

[8] Kuhn, H.W. and R.E. Kuenne, "An Efficient Algorithm for the Numerical Solution of the Generalized Weber Problem in Spatial Economics," Journal of Regional Science, *4*, 21-33 (1962).

[9] Love, R.F., "Locating Facilities in Three-Dimensional Space by Convex Programming," Naval Research Logistics Quarterly, *16*, 503-516 (1969).

[10] Miehle, W., "Link-Length Minimization in Networks," Operations Research, *6*, 232-243 (1958).

[11] Ostresh, L.M., "The Multifacility Location Problem: Applications and Descent Theorems," *Journal of Regional Science, 17*, 409-419 (1977).

[12] Vergin, R.C. and J.D. Rogers, "An Algorithm and Computational Procedure for Locating Economic Activities," Management Science, *13*, 240-254 (1967).

[13] Weiszfeld, E. "Sur le Point pour Lequel la Somme des Distances de n Points Donnes Est Minimum," Tohoku Mathematical Journal, *43*, 355-386 (1936).

AN EASY SOLUTION FOR A SPECIAL CLASS OF FIXED CHARGE PROBLEMS

Patrick G. McKeown

College of Business Administration
University of Georgia
Athens, Georgia

Prabhakant Sinha

Graduate School of Management
Rutgers—The State University
Newark, N.J.

ABSTRACT

The fixed charge problem is a mixed integer mathematical programming problem which has proved difficult to solve in the past. In this paper we look at a special case of that problem and show that this case can be solved by formulating it as a set-covering problem. We then use a branch-and-bound integer programming code to solve test fixed charge problems using the set-covering formulation. Even without a special purpose set-covering algorithm, the results from this solution procedure are dramatically better than those obtained using other solution procedures.

1. INTRODUCTION

The linear fixed charge problem may be formulated as:

(1) Min $\sum_{j \in J} c_j x_j + \sum_{j \in J} f_j y_j$

(2) Subject to $\sum_{j \in J} a_{ij} x_j \geq b_i \quad i \in I,$

(F)

(3) $$y_j = \begin{cases} 1 \text{ if } x_j > 0 \\ 0 \text{ otherwise} \end{cases} \quad j \in J,$$

(4) and $x_j \geq 0, \ j \in J.$

for $I = \{1, \ldots, m\}$ and $J = \{1, \ldots, n\}$.

In addition to continuous costs, the variables have fixed costs which are incurred when the corresponding continuous variable becomes positive. All cost are assumed to be nonnegative. Problem (F) is very similar to the standard linear programming problem, differing only in the presence of the fixed costs. In spite of this similarity, it has proven to be an extremely difficult problem to solve.

If all the continuous costs are zero, we have a special case of the fixed charge problem which we will refer to as problem (PF). Problems of this type can occur, for example, whenever there is a need to find solutions with the least number of basic, nondegenerate variables

In a network context, Kuhn and Baumol [4] discuss the need to know the least number of arcs necessary to carry a desired flow. Also, in the survey processing field, it often becomes necessary to check a record of replies to a questionnaire and to determine changes to make the record consistent. In this case, it is necessary to know the minimum number of such changes that are necessary for consistency. Both of these problems are examples of problem (PF) with the former having the standard transportation constraint matrix and the latter having a general constraint matrix which depends upon the consistency conditions.

A special case of problem (PF) occurs when all the constraint coefficients are nonnegative, i.e., $a_{ij} \geq 0$ for all i, j. We will refer to this problem as (PF+) since we retain the condition that all continuous costs are equal to zero. In this paper, we will demonstrate a solution procedure for (PF+) based on a revised formulation for the problem. We then use a branch-and-bound integer programming code to solve the revised formulation. The results from this approach will be compared to those obtained using other procedures.

2. A REVISED FORMULATION

The problem in which we are interested may be formulated as follows:

(5) Min $\sum_{j \in J} f_j y_j$

(PF+)

subject to (2) – (4)

(6) where $a_{ij} \geq 0$ for $i \in I, j \in J$

(PF+) remains a special case of the fixed charge problem (F) so any results that are applicable to problem (F) will also be applicable to (PF+).

Two previously derived results for (F) that are of particular interest to (PF+) are:

1) any optimal solution to (PF+) will occur at a vertex of the continuous constraint set (2) and (4) (Hirsh and Dantzig [3]);

2) a lower bound, L_0, on the sum of the fixed costs can be found by solving the set-covering problem, P_δ, below (McKeown [5]).

Min $L_0 = \sum_{j \in J} f_j y_j$

(P_δ)

(7) Subject to $\sum_{j \in J} \delta_{ij} y_j \geq 1, \quad i \in I$

(8) $y_j \in (0, 1), \quad j \in J$

(9) where $\delta_{ij} = \begin{cases} 1 \text{ if } a_{ij} > 0 \\ 0 \text{ otherwise,} \end{cases}$

We will combine these two results to develop a solution procedure, the essence of which is summarized in Theorem 1 below.

THEOREM 1: Let $B_\delta^* = \{j | y_j = 1$ in an optimal solution to $P_\delta\}$, then there exists a feasible solution to (PF+) such that $x_j > 0$ for $j \in B_\delta^*$. Furthermore, this solution will be optimal for (PF+).

PROOF: Given an optimal solution to P_δ, we must show that there exists a corresponding solution to (PF+). The first thing to note is that each column of the constraint matrix (7) of P_δ in B_δ^* has at least one nonzero element that is the only nonzero element in that row. Otherwise, the set would be over-covered and we could reduce the objective value of P_δ by removing that column from the optimal solution. We may use this result together with the nonnegativity of the a_{ij} elements to construct a solution to (PF+) using B_δ^*.

Assume, without loss of generality, that $|B_\delta^*| = k$ and that the decision variables have been reindexed such that $\{1, \ldots, k\} \in B_\delta^*$, i.e., the first k variables of (PF+) correspond to the optimal basic variables of P_δ. We can now construct a feasible solution to (PF+) using the following two rules:

$$1) \quad x_1 = \operatorname{Max}_{\substack{a_{i1} \neq 0 \\ i \in I}} \{b_i/a_{i1}\}$$

and

$$2) \quad x_k = \operatorname{Max}\left\{0, \operatorname{Max}_{\substack{a_{ik} \neq 0 \\ i \in I}} \frac{\left\{b_i - \sum_{j=1}^{k-1} a_{ij}x_j\right\}}{a_{ik}}\right\}, \quad k \neq 1$$

This proves the existence of a solution to (PF+) corresponding to B_δ^*. The optimality of this solution is guaranteed by the fact that both (P_δ) and (PF+) have the same objective value and that this objective value for P_δ is a lower bound on (PF+). Hence, B_δ^* corresponds to an optimal solution to P_δ.

3. COMPUTATIONAL COMPARISONS

Since the optimal set of variables for (PF+) can be found by solving the set-covering problem, P_δ, we should be able to use this result to reach quicker solutions to (PF+). We used a mixed integer programming code based on the approach of Tomlin [7] as extended by Armstrong and Sinha [1] to solve the set-covering problems. Special-purpose set-covering algorithms can be expected to perform even better. Fixed charge test problems first generated by Cooper and Drebes [2] were used as a basis of comparison between this set-covering approach and two other procedures. The first such procedure is a branch-and-bound code developed by McKeown [6] specifically for fixed charge problems while the second procedure used the same mixed integer code as before, but solved (PF+) as a mixed integer problem.

The original test problems were of dimension 5×10, but larger problems were generated by putting these smaller problems on the diagonal. Using these problems, the results of our comparisons are shown in Table 1 below.

TABLE I

Problem Set	Size ($m \times n$)	Number of Problems	Average Solution Time per Problem in CPU Seconds on CDC 70/74			
			Armstrong and Sinha	McKeown	Set Covering	LP Solutions
1	5 × 10	12	0.132	0.049	0.017	10
2	10 × 20	6	0.856	0.345	0.046	4
3	15 × 30	4	3.039	1.357	0.101	2

From the table we can see that the set covering formulation is almost three times faster than the best alternative approach for the small (5 × 10) problems and up to 13 times faster for the larger problems (15 × 30). We have also noted the number of problems for which the linear programming solution was integer feasible for the set covering problems. This occurred in over half of the cases.

4. CONCLUSIONS

In this paper, we have shown that a fixed charge problem with nonnegative constraint matrix coefficients and all continuous costs equal to zero can be solved by solving a related set-covering problem. Computational experience confirms that this procedure yields dramatically better solution times than any other available solution procedure. Even quicker solution times can be expected to result if special purpose set-covering codes are used.

REFERENCES

[1] Armstrong, R.D. and P. Sinha, "Improved Penalty Calculations for a Mixed Integer Branch-and-Bound Algorithm," Mathematical Programming, *21*, 474-482 (1974).

[2] Cooper, L. and C. Drebes, "An Approximate Solution Method for the Fixed Charge Problem," Naval Research Logistics Quarterly, *8*, 101-113 (1976).

[3] Hirsch, W.M. and G.B. Dantzig, "The Fixed Charge Problem," Naval Research Logistics Quarterly, *15*, 413-424 (1968).

[4] Kuhn, H.W. and W.J. Baumol, "An Approximative Algorithm for the Fixed-Charges Transportation Problem," Naval Research Logistics Quarterly, *9*, 1-15 (1962).

[5] McKeown, P.G., "A Vertex Ranking Procedure for Solving the Linear Fixed-Charge Problem," Operations Research, *23*, 1183-1191 (1975).

[6] McKeown, P.G., "A Branch-and-Bound Algorithm for the Linear Fixed Charge Problem," Working Paper, University of Georgia (1978).

[7] Tomlin, J.A., "Branch and Bound Methods for Integer and Non-Convex Programming," *Integer and Nonlinear Programming*, 437-450, J. Abadie, Editor, (American Elsevier Publishing Company, New York, 1970).

THE BOUNDED INTERVAL GENERALIZED ASSIGNMENT MODEL

G. Terry Ross

University of Georgia
Athens, Georgia

Richard M. Soland

The George Washington University
Washington, D.C.

Andris A. Zoltners

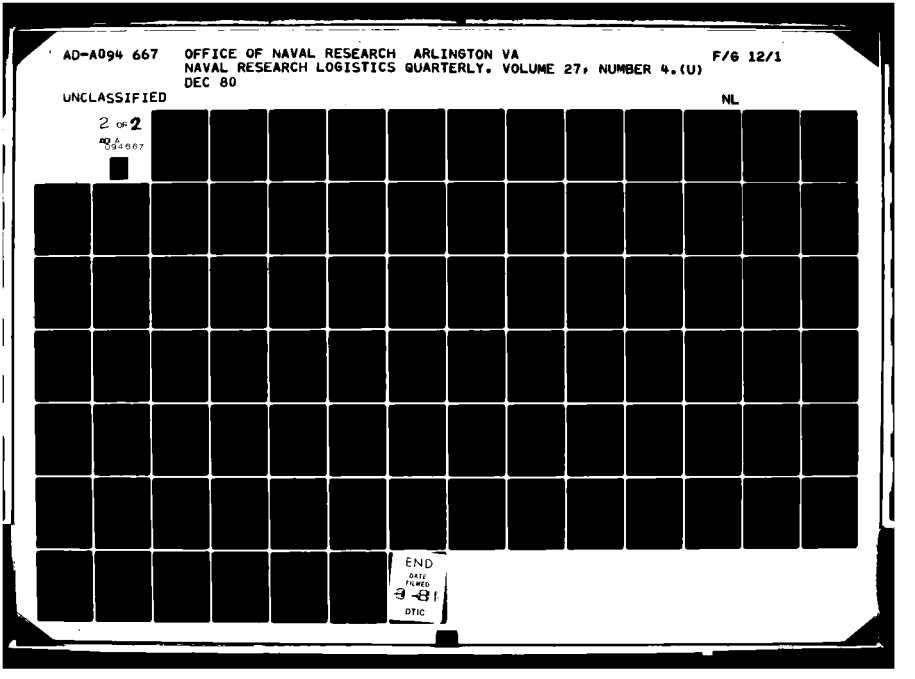
Northwestern University
Evanston, Illinois

ABSTRACT

The bounded interval generalized assignment model is a "many-for-one" assignment model. Each task must be assigned to exactly one agent, however, each agent can be assigned multiple tasks as long as the agent resource consumed by performing the assigned tasks falls within a specified interval. The bounded interval generalized assignment model is formulated, and an algorithm for its solution is developed. Algorithms for the bounded interval versions of the semiassignment model and sources-to-uses transportation model are also discussed.

1. INTRODUCTION

In general terms, assignment models represent problems in which indivisible tasks are to be paired with agents. Given a measure of utility (or disutility) associated with each possible pairing, the objective of the model is to optimize the collective utility associated with assigning a set of tasks to a set of agents. In practical applications, the number of tasks typically exceeds the number of agents, and at least one agent must be assigned two or more tasks if all tasks are to be completed. Examples of such "many-tasks-for-one-agent" problems include the assignment of engagements to a firm's personnel [20], points of distribution to facilities [15], geographic units to district centers [21], products to plants [1], inventory items to warehouses [8], harvestable forest compartments to a labor force [12], ships to shipyards [11], scholarships to students [18], storage compartments to commodities [19], jobs to computers [3], files to mass storage devices [2,13], defect checkpoints to inspectors [17], and trips to ships [7]. The feasibility of many-for-one assignments will depend on the agents' abilities to complete the collections of tasks assigned to them. That is, the subsets of tasks that can be assigned to each agent are determined by the total amount of effort available to the agent and the amount of effort that each individual task requires.


AD-A094 667
OFFICE OF NAVAL RESEARCH ARLINGTON VA
NAVAL RESEARCH LOGISTICS QUARTERLY. VOLUME 27, NUMBER 4.(U)
DEC 80
F/G 12/1
UNCLASSIFIED
NL
2 OF 2
AD A 094667
END
DATE FILMED
4-81
DTIC

Several many-for-one assignment models have been developed which take into account only upper limits on the total amount of effort that each agent may expand. Each of these models is a special case of a model developed by Balachandran [3] and Ross and Soland [14] called the generalized assignment model. This model has the form:

$$(1)\quad (P)\quad \text{minimize}\quad z = \sum_{i \in I} \sum_{j \in J} c_{ij} x_{ij}$$

$$(2)\quad \text{subject to}\quad \sum_{i \in I} x_{ij} = 1 \qquad \text{for all } j \in J,$$

$$(3)\quad \sum_{j \in J} r_{ij} x_{ij} \leqslant b_i \qquad \text{for all } i \in I,$$

$$(4)\quad x_{ij} = 0 \text{ or } 1 \qquad \text{for all } i \in I,\ j \in J.$$

where $I = \{1, 2, \ldots, m\}$ is an agent index set, $J = \{1, 2, \ldots, n\}$ is a task index set, c_{ij} represents the disutility associated with an agent i, task j assignment, $r_{ij} > 0$ denotes the resource burden incurred by agent i in completing task j, and b_i is the resource available to agent i. The decision variable x_{ij} is interpreted as

$$x_{ij} = \begin{cases} 1 \text{ if agent } i \text{ performs task } j \\ 0 \text{ otherwise} \end{cases}$$

Constraints (2) and (4) insure that each task is uniquely assigned to a single agent, and constraints (3) insure that each agent expends no more than b_i resource units in accomplishing assigned tasks. Differences in the difficulty of tasks and differences in agents' abilities to perform the tasks are reflected in the values of the parameter r_{ij}.

The special cases of (P) place various restrictions on the form of the agent resource constraint (3). Francis and White [9] and Barr, Glover and Klingman [5] have addressed the problem in which constraints (3) have the form:

$$(3a)\qquad \sum_{j} x_{ij} \leqslant b_i \quad \text{for all } i \in I.$$

Here b_i denotes the number of jobs agent i can complete, for all jobs consume only one unit of an agent's resource when the agent performs the task (i.e., $r_{ij} = 1$ for all $i \in I$, $j \in J$). The model (1,2,3a,4) is a generalization of the standard assignment problem of linear programming in that it permits an agent to undertake more than one task. It has been called the generalized assignment problem by Francis and White and the semi-assignment problem by Barr, Glover, and Klingman.

Caswell [6], DeMaio and Roveda [8], and Srinivasan and Thompson [16] studied the problem in which (3) is replaced by:

$$(3b)\qquad \sum_{j \in J} r_j x_{ij} \leqslant b_i \quad \text{for all } i \in I.$$

The model (1,2,3b,4) explicitly considers differences in the difficulty of tasks incorporated in the parameter r_j. Srinivasan and Thompson called this model the sources-to-uses problem to reflect the interpretation of the model as a transportation problem in which the demand at the j-th location, r_j, is to be supplied by a single source.

Practical considerations frequently require that the agents expend a minimum total amount of effort in completing assigned tasks. Placing both minimum and maximum restrictions on the resources each agent can expend, yield assignments which neither overburden nor

underutilize the agents. Such restrictions arise in most personnel planning applications [20]. Managerial policies usually require an equitable distribution of work across agents. Analagous restrictions crop up in other contexts as well. For example, in machine loading models, it usually is desirable to balance machine workloads rather than allowing some heavily loaded and some lightly loaded machines. In facility location models, capacity constraints may restrict both the minimum and maximum size of a facility to avoid diseconomies of scale associated with plant sizes outside of a reasonable range, to permit piecewise linear approximation of concave cost functions, or to restrict both the minimum and maximum number of facilities [15]. Similarly, territory design procedures for problems of political districting, school districting, and sales districting require an equitable distribution of some entity (such as voters, minority students, or sales potential) among the districts. Finally, in some applications, upper limits on the effort an agent can expend may be irrelevant, and only lower limits need be considered. Such a situation arises in the segregated storage problem [19] which requires only that a minimal amount of storage space be allocated to store commodities and no maximum allocation is specified.

Thus, from the standpoint of modeling flexibility, it is desirable that assignment models consider explicitly upper and/or lower bounds on the efforts agents must expend in completing assigned tasks. While most "many-for-one" assignment models consider upper bounds, lower bounds have largely been overlooked. In this paper, we introduce the bounded interval generalized assignment model and discuss how existing algorithms can be modified to accommodate lower bounds on agent workloads for this model and its special cases.

2. THE BOUNDED INTERVAL GENERALIZED ASSIGNMENT MODEL AND ALGORITHMIC CONSIDERATIONS

The bounded interval generalized assignment model may be formulated as follows:

$$(5)\quad (P^*)\quad \text{minimize} \quad z = \sum_{i \in I} \sum_{j \in J} c_{ij} x_{ij}$$

$$(6)\quad \text{subject to} \quad \sum_{i \in I} x_{ij} = 1 \qquad \text{for all } j \in J,$$

$$(7)\quad a_i \leqslant \sum_{j \in J} r_{ij} x_{ij} \leqslant b_i \qquad \text{for all } i \in I,$$

$$(8)\quad x_{ij} = 0 \text{ or } 1 \qquad \text{for all } i \in I,\ j \in J.$$

Notice that (P^*) derives from (P). Fortunately, the modeling flexibility achieved through the introduction of lower bounds $a_i > 0$ in constraints (3), (3a), or (3b) does not complicate significantly the computational effort required to solve any of the models described above. Rather, as we shall show, straightforward modifications can be made to the existing algorithms for the semi-assignment problem, sources-to-uses transportation problem, and the generalized assignment problem. The interested reader should consult the cited references for the details of the original algorithms.

In the case of the semi-assignment problem, the constraint matrix is totally unimodular, and integer solutions can be obtained using the simplex method. To impose the lower limit,

$$(7a)\quad \sum_{j} x_{ij} \geqslant a_i \quad \text{for all } i \in I,$$

one need only add an upper bounded slack variable $s_i \leqslant b_i - a_i$ to each of the constraints (3a) and rewrite them as equality constraints. Optimal solutions to the resultant bounded variable linear program will be integer valued.

Models with constraints (3) or (3b) are not totally unimodular. Hence, the solutions of the linear programming relaxation (i.e., $x_{ij} \geqslant 0$ for all i,j) need not be integer. Branch and bound approaches have been developed for deriving integer optimal solutions which solve linear programming relaxations for fathoming and to compute lower bounds. In the case of (3b), a linear programming relaxation is the standard transportation problem [16]; and in the case of (3), a linear programming relaxation is the generalized transportation problem [3]. As in the case of the semi-assignment problem, to impose constraints (7) or

$$(7b) \qquad a_j \leqslant \sum_{j \in J} r_j x_{ij} \leqslant b_i \quad \text{for all } i \in I$$

in a linear programming relaxation, one need only add upper bounded slack variables $s_i \leqslant b_i - a_i$ to constraints (3) or (3b) and rewrite them as equality constraints.

The algorithm developed by Ross and Soland [14] for the generalized assignment problem does not solve a linear programming relaxation to determine the lower bounds. Instead, a Lagrangian relaxation is solved in the form of a series of separable binary knapsack problems. The Lagrangian relaxation has the form:

$$(8) \quad (P_\lambda) \quad \text{minimize} \quad Z_\lambda = \sum_{i \in I} \sum_{j \in J} c_{ij} x_{ij} + \sum_{j \in J} \lambda_j \Big(1 - \sum_{i \in I} x_{ij}\Big)$$

$$\text{subject to} \quad \sum_{j \in J} r_{ij} x_{ij} \leqslant b_i \quad \text{for all } i \in I$$

$$x_{ij} = 0 \text{ or } 1 \text{ for all } i \in I,\ j \in J.$$

The value of each λ_j is set equal to c_{2j}, the second smallest value of c_{ij} for all $i \in I$. These λ_j are optimal dual multipliers for the problem:

$$(P_L) \quad \text{minimize} \quad \sum_{i \in I} \sum_{j \in J} c_{ij} x_{ij}$$

$$\text{subject to} \quad \sum_{i \in I} x_{ij} = 1 \qquad \text{for all } j \in J,$$

$$0 \leqslant x_{ij} \leqslant 1 \qquad \text{for all } i \in I,\ j \in J.$$

Thus, determining a lower bound requires two steps. First, solve (P_L), then solve (P_λ). If the primal solution $\bar{X} = (\bar{x}_{ij})$ to (P_L) should also satisfy (8), then $Z = Z_L = Z_\lambda$, and (P_λ) need not be solved. Frequently, $\bar{X}$ will not satisfy (8), and (P_λ) must be solved to find Z_λ.

To incorporate the lower bounds a_i into the algorithm, one need only replace constraints (8) by constraints (7) giving rise to the problem (P_λ^*) with knapsack constraints bounded both from below and from above. Seemingly, this minor modification to the form of (P_λ) should have little effect on the algorithm. However, it must be noted that (P_λ) will involve fewer 0-1 variables and may be easier to solve than (P_λ^*). The reason is best explained by considering an equivalent form of the objective function of (P_λ):

$$Z_\lambda = \sum_{j \in J} \lambda_j - \text{maximum} \left[\sum_{i \in I} \sum_{j \in J} (\lambda_j - c_{ij}) x_{ij}\right].$$

Clearly, with constraints (8), one can set any x_{ij} equal to zero which has an objective function coefficient $(\lambda_j - c_{ij}) \leq 0$. Thus, using the values of λ_j calculated from solving (P_L), (P_λ) reduces to a problem involving at most $n\,0 - 1$ variables. Such a reduction is not possible for (P_λ^*).

In addition to providing a lower bound, the solutions to (P_L) and (P_λ^*) may be used to select a branching (or separation) variable for defining subsequent candidate problems. As noted above, the solution to (P_L), $\bar{X}$, is usually not feasible to (7). In essence, the solution to (P_λ^*), $\bar{\bar{X}} = (\bar{\bar{x}}_{ij})$, may be interpreted as recommending changes in $\bar{X}$ which must be made in order to satisfy (7). That is, it is possible that for some $j \in J$, $\sum_{i \in I} \bar{\bar{x}}_{ij} = 0$ to avoid overloading any agent or $\sum_{i \in I} \bar{\bar{x}}_{ij} > 1$ to insure every agent uses a minimum amount of his resource. Those variables $\bar{\bar{x}}_{ij}$ with an optimal value of one indicate agent-task pairings that should be made; whereas, those $\bar{\bar{x}}_{ij}$ with an optimal value of zero indicate pairings that should be avoided. Thus, these variable values indicate changes that will reduce the aggregate infeasibility of $\bar{X}$ in (7), and they are helpful in choosing a branching variable.

To formalize the concept of reducing aggregate infeasibility, we define the infeasibility in constraint i prior to taking a branch to be

$$D_i = \max\{0,\ d_i^+,\ d_i^-\}$$

where $$d_i^+ = \sum_{j \in J} r_{ij}\bar{x}_{ij} - b_i,$$

$$d_i^- = a_i - \sum_{j \in J} r_{ij}\bar{x}_{ij}.$$

The set $I^+ \equiv \{i \in I \mid d_i^+ > 0\}$ identifies those constraints (7) for which $\bar{X}$ exceeds the upper bound, and $I^- \equiv \{i \in I \mid d_i^- > 0\}$ identifies those constraints (7) for which $\bar{X}$ fails to satisfy the lower bound.

Suppose $I^+ \neq \emptyset$ and $k \in \{j \in J \mid \bar{x}_{ij} = 1 \text{ and } i \in I^+\}$; if x_{ik} is set to 0 then d_i^+ and d_i^- become:

$$d_i^+ = \sum_{j \in J} r_{ij}\bar{x}_{ij} - b_i - r_{ik}$$

$$d_i^- = a_i - \sum_{j \in J} r_{ij}\bar{x}_{ij} + r_{ik}$$

and the resulting infeasibility in constraint i becomes

$$D_i^k = \max\{0,\ d_i^+,\ d_i^-\}.$$

Assuming that task k is reassigned to the second least costly agent, (say agent h, where $c_{hk} = \min_{l \neq i} c_{lk}$) then the infeasibility in constraint h becomes

$$D_h^k = \max\{0,\ d_h^+,\ d_h^-\}$$

where

$$d_h^+ = \sum_{j \in J} r_{hj}\bar{x}_{hj} - b_h + r_{hk}$$

$$d_h^- = a_h - \sum_{j \in J} r_{hj}\bar{x}_{hj} - r_{hk}.$$

Hence, the net difference in total infeasibility is:

$$\Delta D_i^k = (D_i + D_h) - (D_i^k + D_h^k)$$

If $\Delta D_i^k > 0$ then setting $x_{ik} = 0$ yields a reduction in aggregate infeasibility, and if $\Delta D_i^k \leqslant 0$ then such a branch will not reduce aggregate infeasibility.

Similarly, suppose that $I^- \neq \emptyset$ and $k \in \{j \in J | \bar{x}_{ij} = 0 \text{ and } i \in I^-\}$; if x_{ik} were set to 1 then d_i^+ and d_i^- become:

$$d_i^+ = \sum_{j \in J} r_{ij} \bar{x}_{ij} - b_i + r_{ik}$$

$$d_i^- = a_i - \sum_{j \in J} r_{ij} \bar{x}_{ij} - r_{ik}$$

and the resulting infeasibility in constraint i would be

$$D_i^k = \max \{0, d_i^+, d_i^-\}.$$

If x_{ik} is set to 1 then task k is assigned to agent i and agent i_k relinquishes it, where $i_k = \min_i c_{ik}$. Hence, the infeasibility for constraint i_k becomes

$$D_{i_k}^k = \max \{0, d_{i_k}^+, d_{i_k}^-\}$$

where

$$d_{i_k}^+ = \sum_{j \in J} r_{i_k j} \bar{x}_{i_k j} \epsilon - b_{i_k} - r_{i_k k}$$

$$d_{i_k}^- = a_{i_k} - \sum_{j \in J} r_{i_k j} \bar{x}_{i_k j} + r_{i_k k}.$$

The net difference in infeasibility is

$$\Delta D_i^k = (D_i + D_{i_k}) - (D_i^k + D_{i_k}^k)$$

where D_i and D_{i_k} are the infeasibilities in constraints i and i_k prior to any branch. As before, if $\Delta D_i^k > 0$ then there is reduced infeasibility following a branch on variable x_{ik}.

Several rules for selecting the branching variable, $x_{i^\bullet j^\bullet}$, are formulated as follows:

1. a) $x_{i^\bullet j^\bullet}$ is that variable for which

$$\Delta D_{i^\bullet}^{j^\bullet} = \max_{(i,j) \in H^+ \cup H^-} \{\Delta D_i^j\}$$

where $H^+ \equiv \{(i, j) | \bar{\bar{x}}_{ij} = 0 \text{ and } i \in I^+\}$

$H^- \equiv \{(i, j) | \bar{\bar{x}}_{ij} = 1 \text{ and } i \in I^-\}$

b) If $\Delta D_{i^\bullet}^{j^\bullet} = 0$ in a) then $x_{i^\bullet j^\bullet}$ is that variable for which

$$\Delta D_{i^\bullet}^{j^\bullet} = \max_{(i,j) \in G^+ \cup H^+ \cup G^- \cup H^-} \{\Delta D_i^j\}$$

where $G^+ \equiv \{(i, j) | \bar{x}_{ij} = 1 \text{ and } i \in I^+\}$

$G^- \equiv \{(i, j) | \bar{x}_{ij} = 0 \text{ and } i \in I^-\}$.

II. a) $x_{i^\bullet j^\bullet}$ is that variable for which

$$\rho_{i^\bullet j^\bullet} = \min\left\{ \min_{(i,j) \in G^+} \left\{ \frac{\lambda_j - c_{ij}}{\Delta D_i^j} \right\}, \min_{(i,j) \in G^-} \left\{ \frac{c_{ij} - c_{ikj}}{\Delta D_i^j} \right\} \right\}$$

b) If $\Delta D_i^j = 0$ for all $(i, j) \in G^+ \cup G^-$ then $x_{i^\bullet j^\bullet}$ is that variable for which

$$\rho_{i^\bullet j^\bullet} = \max_{(i,j) \in E^+} \left\{ \frac{\lambda_j - c_{ij}}{\left\lceil r_{ij} / \left(b_i - \sum_{j \in F_i} r_{ij}\right) \right\rceil} \right\}$$

where $E^+ \equiv \{(i, j) | \bar{x}_{ij} = 1 \text{ and } i \in I^+\}$,

F_i denotes the set of tasks assigned to agent i by prior branching.

Rules Ia and Ib are designed to choose that variable which reduces the post branch aggregate infeasibility by the greatest amount. Rule IIa conditions the choice of branching variable on the additional cost incurred per unit reduction in infeasibility. Rule IIb is the one used in [14]; the variable chosen by this rule represents an agent-task pairing which should be made considering the penalty for not doing so weighted by the fraction of the agent's remaining free resources consumed by the assignment.

As the algorithm progresses and new candidate problems (CPs) are defined by the branching process, the additional steps given below may be taken to facilitate fathoming. These steps are specialized adaptations of more general forcing (or variable fixing) tests suggested by Balas [4] and Glover [10].

In solving any (CP), any $x_{i'j'}$ for which $r_{i'j'} > b_{i'} - \sum_{j \in F_{i'}} x_{i'j} r_{i'j}$ may be set equal to zero. Here $F_{i'}$ denotes those $j \in J$ for which $x_{i'j}$ has been assigned a value of zero or one by prior branching or variable fixing tests. Similarly, if there is an $x_{i'j'}$ for which $a_{i'} - \sum_{j \in F_{i'}} x_{i'j} r_{i'j} > \sum_{j \in J - F_{i'}} r_{i'j} - r_{i'j'}$, then $x_{i'j'}$ must be set equal to one in the solution to (CP). These variable forcing tests may subsequently result in other variables being forced to zero or to one when all of the resultant implications are considered. Moreover, forcing certain variables to zero or to one in the solution to (CP) may affect the values of some of the λ_j obtained from solving (P_I^*). This change may, in turn, increase the value of the lower bound provided by (P_λ^*).

Another test may be used to check the feasibility of (P^*) (or any candidate subproblem). Summing the constraints (7) together yields the constraint (9):

$$A = \sum_{i \in I} a_i \leqslant \sum_{i \in I} \sum_{j \in J} r_{ij} x_{ij} \leqslant \sum_{i \in I} b_i = B. \tag{9}$$

This new constraint, together with constraints (6), implies that for any feasible solution to (P^*) we must have:

$$\sum_{j \in J} r_j' \geqslant A \text{ and } \sum_{j \in J} r_j'' \leqslant B \tag{10}$$

where

$$r_i' \equiv \max_{i \in I} \{r_{ij}\} \text{ and } r_i'' \equiv \min_{i \in I} \{r_{ij}\}.$$

The values necessary for the tests (10) can be updated easily as part of the branching process in order to apply this test to each (CP).

The algorithm terminates in the usual way when all candidate problems have been fathomed.

3. CONCLUSION

This note has described an efficient branch and bound algorithm for the bounded interval generalized assignment problem. The algorithm serves as a useful tool for solving a large number of applications of this assignment model, a representative sample of which is mentioned in the introduction.

REFERENCES

[1] Abella, R.J. and T.E. Bova, "Optimal Plant Allocation of Stockkeeping Units," presented at TIMS/ORSA Joint National Meeting, San Francisco, California (May 1977).

[2] Babad, J.M., V. Balachandran and E.A. Stohr, "Management of Program Storage in Computers," Management Science, *23*, 380-390 (1976).

[3] Balachandran, V., "An Integer Generalized Transportation Model for Optimal Job Assignment in Computer Networks," Operations Research, *24*, 742-759 (1976).

[4] Balas, E., "*An Additive Algorithm for Solving Linear Programs with Zero-one Variables*," Operations Research, *13*, 517-545 (1965).

[5] Barr, R.S., F. Glover and D. Klingman, "A New Alternating Basis Algorithm for Semi-assignment Networks," Research Report CCS-264, Center For Cybernetic Studies, University of Texas, Austin, Texas (January 1977).

[6] Caswell, W., "The Transignment Problem," Unpublished Ph.D. Thesis, Rensselaer Polytechnic Institute (1972).

[7] Debanne, J.G. and J-N Lavier, "Management Science in the Public Sector—The Estevan Case," Interfaces, *9*, 66-77 (1979).

[8] DeMaio, A. and C. Roveda, "An All Zero-One Algorithm for a Certain Class of Transportation Problems," Operations Research, *19*, 1406-1418 (1971).

[9] Francis, R.L. and J.A. White, *Facility Layout and Location: An Analytical Approach*, (Prentice-Hall, Englewood Cliffs, New Jersey, 1974).

[10] Glover, F., "A Multiphase-Dual Algorithm for the Zero-One Integer Programming Problem," Operations Research, *13*, 879-919 (1965).

[11] Gross, D. and C.E. Pinkus, "Optimal Allocation of Ships to Yards for Regular Overhauls," Technical Memorandum 63095, Institute for Management Science and Engineering, The George Washington University, Washington, D.C. (May 1972).

[12] Littschwager, J.M. and T.H. Tcheng, "Solution of a Large-scale Forest Scheduling Problem by Linear Programming Decomposition," Journal of Forestry, *65*, 644-646 (1967).

[13] Morgan, H.L., "Optimal Space Allocation on Disk Storage Devices," Communications of the ACM, *17*, 139-142 (1974).

[14] Ross, G.T. and R.M. Soland, "A Branch and Bound Algorithm for the Generalized Assignment Problem," Mathematical Programming, *8*, 91-103 (1975).

[15] Ross, G.T. and R.M. Soland, "Modeling Facility Location Problems as Generalized Assignment Problems," Management Science, *24*, 345-357 (1977).

[16] Srinivasan, V. and G.L. Thompson, "An Algorithm For Assigning Uses to Sources in a Special Class of Transportation Problems," Operations Research, *21*, 284-295 (1973).

[17] Trippi, R.R, "The Warehouse Location Formulation as a Special Type of Inspection Problem," Management Science, *21*, 986-988 (1975).

[18] Wagner, H.M., *Principles of Operations Research*, (Prentice-Hall, Englewood Cliffs, N.J., 1968).

[19] White, J.A. and R.L. Francis, "Solving A Segregated Storage Problem Using Branch and Bound and Extreme Point Ranking," AIIE Transactions, *3*, 37-44 (1971).

[20] Zoltners, A.A., "The Audit Staff Assignment Problem: An Integer Programming Approach," Working Paper 74-34, School of Business Administration, University of Massachusetts, Amherst, Massachusetts (September 1974).

[21] Zoltners, A.A., "A Unified Approach to Sales Territory Alignment," *Sales Management: New Developments from Behavioral and Decision Model Research* R. Bagozzi, Editor, (Cambridge, Massachusetts Marketing Science Institute, 1979), 360-376.

THE M/G/1 QUEUE WITH INSTANTANEOUS BERNOULLI FEEDBACK*

Ralph L. Disney

Virginia Polytechnic Institute and State University
Blacksburg, Virginia

Donald C. McNickle

University of Canterbury
Christchurch, New Zealand

Burton Simon

Bell Laboratories
Holmdel, New Jersey

ABSTRACT

In this paper we are concerned with several random processes that occur in *M/G/1* queues with instantaneous feedback in which the feedback decision process is a Bernoulli process. Queue length processes embedded at various times are studied. It is shown that these do not all have the same asymptotic distribution, and that in general none of the output, input, or feedback processes is renewal. These results have implications in the application of certain decomposition results to queueing networks.

1. INTRODUCTION

In this paper we are concerned with several random processes that occur within the class of *M/G/1* queues with instantaneous feedback in which the feedback decision process is a Bernoulli process. Such systems in the case $G = M$ are among the simplest, nontrivial examples of Jackson networks [8]. Indeed, they are so simple that they are usually dismissed from consideration in queueing network theory as being obvious. We will show that far from being obvious, they exhibit some important unexpected properties whose implications raise some interesting questions about Jackson networks and their application.

In particular, Jackson [8] observed that in his networks the vector-valued queue length process behaved as if the component processes were independent, *M/M/1* systems. Since those results appeared there has developed a mythology to explain them. These arguments usually rest on three sets of results that are well known in random point process theory: superposition, thinning, and stretching. By examining the network flow, it will be shown that the applications of these results are inappropriate for queueing networks with instantaneous, Bernoulli feedback. These flows are considerably more complicated than one expects based on such arguments.

*The research was supported under ONR Contracts N00014-75-C-0492 (NR042-296) and N00014-77-C-0743 (NR042-296)

It is shown that in general, both the input and output processes of the *M/M/1* queue with feedback are Markov-renewal, and the kernels of those Markov-renewal processes are given. The output of the *M/G/1* queue with feedback is also Markov-renewal, and that kernel is given. It is shown that in general these processes are never renewal. The implications of these facts are discussed in Section 4.

1.1 The Problem and Notation

We assume the usual apparatus of an *M/G/1* queue with unlimited waiting capacity. The new idea is that a unit which has received service departs with probability q and returns for more service with probability p, $p + q = 1$. Without loss of generality for the processes studied here, the returning customer can be put anywhere in the queue.

To establish notation it is assumed that the *arrival process* is a Poisson process with parameter $\lambda > 0$. The arrival epochs are the elements of $\{W_n: n = 1, 2, \ldots\}$. Service times are independent, identically distributed, nonnegative, random variables, S_n with

$$Pr\{S_n \leqslant t\} = H(t), \quad t \geqslant 0,$$
$$E\{S_n\} < \infty.$$

We define $H^*(s)$, the Laplace-Stieltjes transform of $H(t)$, by

$$H^*(s) = \int_0^\infty e^{-st} dH(t), \quad \text{Re } s \geqslant 0.$$

The arrival process and service times are independent processes.

Service completions occur at $T_0 < T_1 < T_2 \ldots$ called the *output epochs*. Let

$$Y_n = Y_{T_n} = \begin{cases} 0, & \text{if the } n\text{-th output departs,} \\ 1, & \text{if the } n\text{-th output feeds back.} \end{cases}$$

$\{Y_n\}$ is a Bernoulli process.

Elements of the subset $\{t_n\} \subset \{T_n\}$ are called the *departure epochs* and are the times at which an output leaves the system. The elements of the subset $\{\tau_n\} \subset \{T_n\}$ are called the *feedback epochs* and are the times at which an output returns to the queue. $\{t_n\} \cup \{\tau_n\} = \{T_n\}$.

The times T_n' are the times at which a unit enters the queue. $\{T_n'\}$ is called the *input process*. $\{T_n'\} = \{W_n\} \cup \{\tau_n\}$.

There are five queue length processes to be studied. They are closely related as will be shown. Let $Q(t)$ be the queue length (number in the system) at t. Then, $Q_1(n) = Q(W_n - 0)$; $Q_2(n) = Q(T_n' - 0)$; $Q_3^*(n) = Q(T_n + 0)$; $Q_4^*(n) = Q(t_n + 0)$ are respectively the embedded queue lengths at arrival epochs, input epochs, output epochs, departure epochs.

2. QUEUE LENGTH PROCESSES

The queue lengths listed in Section 1.1 are closely related. The steady state versions of $\{Q_3^*(n)\}$ and $\{Q_4^*(n)\}$ are of primary concern. They are studied in Sections 2.1 and 2.2 separately. They are related to the other processes in Section 2.3. The important special case for $G = M$ is then studied in 2.4.

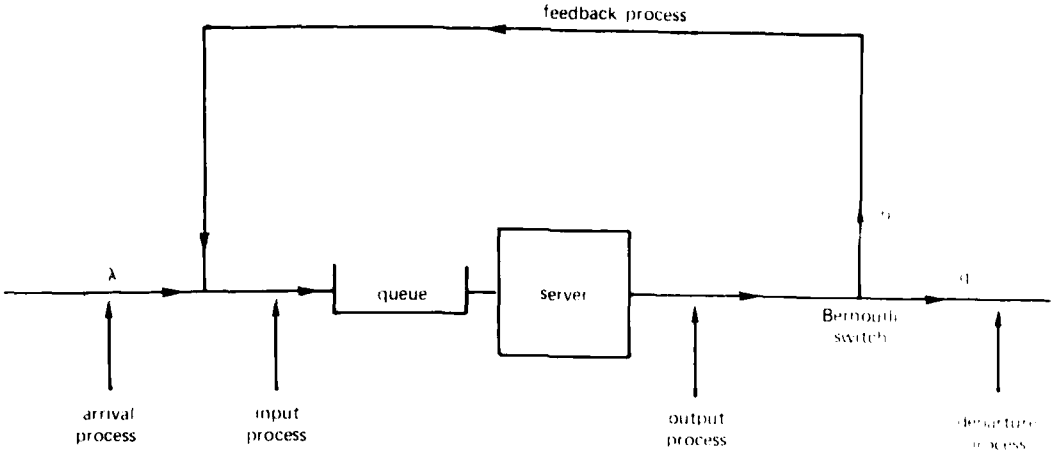


FIGURE 1.

2.1 The $\{Q_4^+(n)\}$ Process

There are several ways to study this process. The following appears to be direct, correct, and may help explain why these feedback problems have received such little attention in the queueing literature. First, it is clear that

$$t_n = \begin{cases} t_{n-1} + S_n', & \text{if } Q_4^+(n-1) > 0, \\ t_{n-1} + I_n + S_n', & \text{if } Q_4^+(n-1) = 0. \end{cases}$$

Here S_n' is the total service time consumed between the $(n-1)-st$ and n-th departure. I_n is the idle time following t_{n-1} when $Q_4^+(n-1) = 0$. For the *M/G/1* queue, the I's are independent, identically distributed, random variables that are exponentially distributed with parameter λ.

Without loss of generality, since customers are indistinguishable,

$$S_n' = S_1 + S_2 + \ldots S_m,$$

where m is the number of services performed between the $(n-1)-st$ and n-th departure. Since $\{Y_n\}$ is a Bernoulli process, m is geometrically distributed and it follows that $\{S_n'\}$ is a sequence of independent, identically distributed, random variables. Thus, the Laplace-Stieltjes transform of the distribution function of S_n' is easily found to be

$$G^*(s) = qH^*(s)/[1 - pH^*(s)].$$

Using standard embedded Markov chain methods [3, 167-174] one finds that the probability generating function of $\prod'$, the limiting probability distribution of $\{Q_4^+(n)\}$, is given by

$$\hat{g}(z) = \frac{\pi'(0)(z-1)G^*(\lambda - \lambda z)}{z - G^*(\lambda - \lambda z)} \tag{1}$$

and

$$\pi'(0) = 1 - \lambda E[S_n]/q. \tag{2}$$

If one is willing to assume that the *M/G/1* queue with instantaneous, Bernoulli feedback has a queue length process which asymptotically has the same distribution as another *M/G/1* queue without feedback, then (1) and (2) follow immediately. This assumption is valid since if customers feedback to the front of the queue, the total service time of the n-th customer is S_n'. $\{S_n'\}$ is a sequence of independent, identically distributed, random variables with mean $E[S_n]/q$. Alternatively, one can argue that the *M/G/1* queue with feedback (as defined here) has the same asymptotic distribution for its queue length process as an *M/G/1* queue without feedback if one takes the arrival process parameter in the latter case to be λ/q. Indeed, both of these assumptions and several others that are used to "prove" that these queues with feedback are trivial have now been proven by the arguments leading up to (1) and (2). That these arguments can be applied more generally is easily proven. In the remainder of this paper, in Takács [10] and in Disney [6] it is shown that while these arguments may imply that the study of queue lengths at departure times is trivial, the same cannot be said for other processes of interest.

2.2 The $\{Q_3^+(n)\}$ Process

This is the queue length process embedded at output points. Since $\{t_n\} \subset \{T_n\}$, $\{Q_4^+(n)\}$ is a process on a coarser grid than $\{Q_3^+(n)\}$. Since one is ultimately to be concerned with both $\{Q_3^+(n)\}$ and $\{T_n - T_{n-1}\}$, the following study is for the joint process $\{Q_3^+(n),\ T_n - T_{n-1}\}$. The marginal results for $\{Q_3^+(n)\}$ then will be easy to determine.

THEOREM 1: The process $\{Q_3^+(n),\ T_n - T_{n-1}\}$ is a Markov-renewal process with kernel $A(i,j,x) = Pr\{Q_3^+(n) = j,\ T_n - T_{n-1} \leqslant x \mid Q_3^+(n-1) = i\}$. If one defines

$$P_j(y) = (\lambda y)^j e^{-\lambda y}/j!, \qquad j = 0, 1, 2, \ldots,$$

then

$$A(i,j,x) = \begin{cases} 0, & \text{if } j < i-1, \\ \int_0^x (P_{i-1}(y)p + P_{i-j-1}(y)q)\,dH(y), & \text{if } i \neq 0, \\ & j \geqslant i-1, \\ \int_0^x (1 - e^{-\lambda(y-t)})\,(P_{j-1}(y)p + P_j(y)q)\,dH(y), & \text{if } i = 0, \\ & j > 0, \\ \int_0^x (1 - e^{-\lambda(y-t)})\,P_0(y)q\,dH(y), & \text{if } j = i = 0. \end{cases}$$

PROOF:

$$T_n - T_{n-1} = \begin{cases} S_n, & \text{if } Q_3^+(n-1) > 0, \\ I_n + S_n, & \text{if } Q_3^+(n-1) = 0, \end{cases}$$

where I_n is the exponentially distributed idle time preceeding S_n if $Q_3^+(n-1) = 0$. The result then follows directly using arguments as in [5]. □

As $x \to \infty$, $A(i,j,x) \to A(i,j)$ the one step transition probability for the $\{Q_3^+(n)\}$ process. Then using standard embedded Markov chain results [3, 167-174] one can show that the probability generating function $g(z)$ for the limiting probabilities $\pi(j)$ of $Q_3^+(n)$ are given by

$$g(z) = \frac{\pi(0)(z-1)(pzH^*(\lambda - \lambda z) + qH^*(\lambda - \lambda z))}{z - pzH^*(\lambda - \lambda z) - qH^*(\lambda - \lambda z)} \tag{3}$$

and

$$\pi(0) = q - \lambda E[S_n]. \tag{4}$$

2.3 Other Queue Length Processes

The queue length and limiting probabilities for the queueing processes, $\{Q_1(n)\}$, $\{Q_2^-(n)\}$ now follow from a theorem found in Cooper [3, 155]. From this it follows that $\{Q(t)\}$, $\{Q_1^-(n)\}$, and $\{Q_4^+(n)\}$ are asymptotically, identically distributed (see Cooper [3, 65]) and $Q_2^-(n)$, and $\{Q_3^+(n)\}$ are asymptotically, identically distributed. Clearly, $\{Q_4^+(n)\}$ and $\{Q_3^+(n)\}$ are not asymptotically, identically distributed. That $\{Q_4^+(n)\}$ and $\{Q_3^+(n)\}$ are not asymptotically, identically distributed can be seen as follows. First, in the set up of studying the $\{Q_3^+(n)\}$ process one must decide how to count the feedback customer when he appears. The clean way to do this is to use Y_n as defined in Section 1.1 and ${}^*Q_3^+(n)$ as the number in the queue not including the outputting customer. Then one can study the process $\{Y_n, {}^*Q_3^+(n)\}$. Indeed, this is precisely the direction used, for example, in d'Avignon and Disney [4]. Then the $\{Q_3^+(n)\}$ of Theorem 1 above would be the $\{{}^*Q_3^+(n) + Y_n\}$ process of [4]. It then follows that $\{{}^*Q_3^+(n)\}$ and $\{Q_4^+(n)\}$ are asymptotically, identically distributed. Thus, if one does not count the feedback customer in the queue length process, the queue length processes defined in Section 1.1 are all asymptoticaly, identically distributed.

2.4 The *M/M/*1 Case

If one assumes that the service time distribution is

$$H(t) = 1 - e^{-\mu t}, \quad t \geqslant 0,$$

some further clarification is possible here. From the results of Jackson [8],

$$\pi'(j) = \left(1 - \frac{\lambda}{q\mu}\right)\left(\frac{\lambda}{q\mu}\right)^j, \quad j = 0, 1, 2, \ldots.$$

From (3) and (4) one obtains

$$\pi(0) = q\left(1 - \frac{\lambda}{q\mu}\right),$$

$$\pi(j) = \left(1 - \frac{\lambda}{q\mu}\right)\left(\frac{\lambda}{q\mu}\right)^{j-1}\left(p + \frac{\lambda}{\mu}\right), \quad j = 1, 2, \ldots.$$

Comments in Section 2.3 explain this difference between $\pi(j)$ and $\pi'(j)$.

3. FLOW PROCESSES

To further clarify the problems here, it is useful to study the flow processes in this system. There are five processes of interest: the arrival process, the input process, the output process, the departure process, and the feedback process.

There have been some questions since the publication of the Jackson results concerning the interpretation of his results [2]. In his paper Jackson showed that for his networks the joint limiting probability for the vector of queue lengths at each server could be factored into limiting probabilities for the queue length at each server. This imples that the queue lengths are independent in the limit. Furthermore, the marginal limiting probabilities were found to be

precisely those of an $M/M/1$ queue. Burke [2], has argued that the Jackson results are surprising. Burke's argument is based on showing that the input to a single server queue with feedback is not Poisson because the interinput times (our $\{T'_n - T'_{n-1}\}$) are not exponentially distributed. [2] gives the precise result

$$Pr\{T'_n - T'_{n-1} \leqslant t\} = 1 - \frac{q\mu - \lambda}{\mu - \lambda} e^{-\lambda t} - \frac{p\mu}{\mu - \lambda} e^{-\mu t}, \quad t \geqslant 0.$$

In this section we will study some of the flows in this network and show indeed that the Jackson results are surprising.

3.1 Departures

The departure process $\{t_n\}$ can be studied as in Disney, Farrell, deMorais [5] upon using the mapping in Section 2.1. Thus we know that whenever $\{S_n\}$ is a renewal process with exponential distribution this departure process is a renewal process, and is a Poisson process. This is the Jackson case. So we conclude that the departure process from the Jackson network is a Poisson process.

From the results of Section 2.1 it would seem possible that the departure process is Poisson even if S_n is not exponentially distributed. The result that is needed for the results of [5] to follow is that S'_n be exponentially distributed (since it is known that $\{S'_n\}$ is a sequence of mutually independent, identically distributed, random variables).

LEMMA 1: The departure process from the $M/G/1$ queue with feedback is a renewal process if and only if S_n is exponentially distributed for every n. In that case the departure process is Poisson.

PROOF: From Section 2.1 we have $G^*(s)$, the Laplace-Stieltjes transform of the distribution functions of S'_n is given by

$$G^*(s) = \frac{qH^*(s)}{1 - pH^*(s)}.$$

From [5], when the queue capacity is infinite the departure process will be a renewal process if and only if S'_n is exponentially distributed with parameter a, and will be Poisson in that case. But this implies that $H^*(s)$ must satisfy

$$a/(a + s) = qH^*(s)/[1 - pH^*(s)].$$

The only solution here is

$$H^*(s) = \frac{a/q}{a/q + s}$$

which implies $H(t)$ is exponential.

3.2 Outputs and Inputs

From Section 2.2 it is clear that the output process is a Markov-renewal process whose distributions are given by $A(i,j,x)$. From these, the following results are obtained.

THEOREM 2: The output process $\{T_n - T_{n-1}\}$ is a renewal process if and only if $q = 1$ and $H(t) = 1 - e^{-\mu t}$.

PROOF: If $q = 1$ and $H(t) = 1 - e^{-\mu t}$, the output process and departure process are identical processes. Furthermore, the processes are both departure processes from a *M/M/1* queue without feedback. From [5] we have that this departure process is a Poisson process and "if" follows. To prove "only if" we consider the contrapositive statement and assume $q \neq 1$. (The other side of the contrapositive would have $H(t) \neq 1 - e^{-\mu t}$. But then "only if" follows trivially from [5]. Thus, we need only consider the case of $q \neq 1$.) Equations (3.1) and (3.2) in [5] can be modified in such a way that one can show that if $q \neq 1$, there is no solution to both of those equations simultaneously. Then using the same arguments as in [5] one has that $\{T_n - T_{n-1}\}$ is not a renewal process and therefore "only if" is proven. □

To be more specific, Theorem 2 can be particularized as

COROLLARY 1: The output process $\{T_n - T_{n-1}\}$ for the *M/M/1* queue is a Poisson process if and only if $q = 1$. One can prove this result (in fact it is obvious) directly from Theorem 2. The following is an alternate proof that exposes a bit more of the properties of these systems. Again we use a contrapositive proof for "only if".

PROOF: Define

$$F(x) = Pr\{T_n - T_{n-1} \leqslant x\}.$$

$F(x) = \pi A U$ where U is a column vector all of whose elements are 1, π is the vector of limiting probabilities given in Section 2.4 for $\{Q_3^+(n)\}$ and A is the matrix of $A(i,j,x)$. Then from Theorem 1 one obtains after some algebraic manipulations:

$$F(x) = \left(q - \frac{\lambda}{\mu}\right) \int_0^x [1 - e^{-\lambda(x-y)}]\, dH(y) + \left(p + \frac{\lambda}{\mu}\right) H(x) \tag{5}$$

for any $M/G/1$ queue with instantaneous, Bernoulli feedback.

For $H(y) = 1 - e^{-\mu y}$, it follows that

$$F(x) = 1 - \frac{q\mu - \lambda}{\mu - \lambda} e^{-\lambda x} - \frac{p\mu}{\mu - \lambda} e^{-\mu x}, \quad x \geqslant 0. \tag{6}$$

Thus, single intervals are not exponentially distributed and the output process is not a Poisson process if $q \neq 1$. On the other hand if $q = 1$, then we fulfill the conditions of Theorem 2. Hence, $\{T_n - T_{n-1}\}$ is a renewal process. But from (6) this renewal process has exponentially distributed intervals and thus is a Poisson process. □

Formula (6) was previously found by Burke [2] for the distribution of times between inputs. The input process can be analyzed as follows:

THEOREM 3: If $H(x) = 1 - e^{-\mu x}$, the process $\{Q_2(n),\ T_n' - T_{n-1}'\}$ is a Markov-renewal process with kernel

$$Y(i,j,x) = Pr[Q_2(n) = j,\ T_n' - T_{n-1}' \leqslant x \mid Q_2(n-1) = i]$$

given by

$$Y(i,j,x) = \begin{cases} 0; & j > i + 1, \\ \displaystyle\int_0^x (e^{-\lambda s} - qe^{-\lambda s})\, q^i\, dH^{(i+1)}(s); & j = 0,\ i \geqslant 0, \\ \displaystyle\int_0^x e^{-\lambda s}\left(\frac{q\lambda}{\lambda + \mu}(1 - e^{-(\lambda+\mu)(x-s)}) + p\right) q^{i-j}\, dH^{(i-j+1)}(s); & 1 \leqslant j \leqslant i, \\ e^{-\mu x}(1 - e^{-\lambda x}); & j = i + 1, \end{cases}$$

where $dH^{(n+1)}(s) = \dfrac{\mu(\mu s)^n e^{-\mu s}}{n!}\,ds$.

PROOF: Clearly, if $j > i+1$ then $Y(i,j,x) = 0$. If $j = 0$ then $Y(i,j,x)$ is the probability that the $i+1$ customers in line all depart before x and the first arrival occurs after the last departure, but before x; or, the first $i-1$ customers depart, but the last one feeds back before x, and there are no arrivals while this is happening.

If $1 \leqslant j \leqslant i$ then $Y(i,j,x)$ is the probability that $i-j+1$ customers depart before x, no arrivals occur during this time, but between the last departure and x, an arrival occurs before a departure; or, $i-j+1$ customers are served before x, the first $i-j$ depart, the last one feeds back, and there are no arrivals while this is happening.

If $j = i+1$ then $Y(i,j,x)$ is the probability that there is an arrival before x and no departures before x. Since $Y(i,j,x)$ never depends on $\{Q_2^-(k);\ k < n-1\}$ or $\{T_k';\ k < n\}$, the process $\{Q_2^-(n),\ T_n' - T_{n-1}'\}$ is a Markov-renewal process. □

Now, if $Y(x)$ is the matrix whose elements are $Y(i,j,x)$, π is the vector of probabilities found in (3) and U is a vector all of whose elements are 1 then it is easy to see that

$$F(x) = Pr[T_n' - T_{n-1}' \leqslant x] = \pi Y(x) U$$

and

$$G(x,y) = Pr[T_n' - T_{n-1}' \leqslant x,\ T_{n+1}' - T_n' \leqslant y] = \pi Y(x) Y(y) U,$$

where $F(x)$ is the $F(x)$ given by (6). Of course, if $\{T_n' - T_{n-1}'\}$ is to be a renewal process then it is necessary (but not sufficient) that

$$G(x,y) = F(x)F(y).$$

From this we can conclude:

COROLLARY 2: The input process to the *M/M/1* queue with instantaneous, Bernoulli feedback is not a renewal process unless $q = 1$.

PROOF: If $q = 1$ then the input process is just the arrival process which is Poisson.

If the input process is a renewal process for $q \neq 1$ then it must be true that

$$\forall x;\ \pi Y(x) U = F(x)$$

$$\forall x,y;\ \pi Y(x) Y(y) U = F(x)F(y) \text{ where}$$

$F(x)$ is given by (6) and U is a column of 1's. Thus,

$$\forall x,y;\ \left[\pi - \frac{\pi Y(x)}{F(x)}\right] Y(y) U = 0.$$

Some algebra yields

$$\left[\pi - \frac{\pi Y(x)}{F(x)}\right] Y(y) U = \frac{F(x) - (1 - e^{-\mu x})}{F(x)} \left[\frac{p\lambda}{\mu - \lambda} e^{-\lambda x} - \frac{p\mu}{\mu - \lambda} e^{-\mu x} + p e^{-(\mu+\lambda)x}\right].$$

If $q \neq 1$,

$$F(x) - (1 - e^{-\mu x}) = e^{-\mu x} - \frac{\mu q - \lambda}{\mu - \lambda} e^{-\lambda x} - \frac{\mu p}{\mu - \lambda} e^{-\mu x}$$

$$< e^{-\mu x} - \frac{\mu q - \lambda}{\mu - \lambda} e^{-\mu x} - \frac{\mu p}{\mu - \lambda} e^{-\mu x} = 0.$$

Thus, to show that the input process is not renewal, it suffices to show that for some y,

$$\frac{p\lambda}{\mu-\lambda}e^{-\lambda y}-\frac{p\mu}{\mu-\lambda}e^{-\mu y}+pe^{-(\mu+\lambda)y}\neq 0.$$

The third term of the Taylor expansion of this expression is

$$\frac{p\lambda}{\mu-\lambda}\frac{\lambda^2}{2}-\frac{p\mu}{\mu-\lambda}\frac{\mu^2}{2}+\frac{p(\mu+\lambda)}{2}=\frac{p\lambda\mu}{2}\neq 0,$$

so (by [1, 198] for instance), it cannot be identically zero unless $p = 0$ (i.e., $q = 1$). □

It seems obvious that the arrival process and feedback process are not independent processes. One can show, using the above arguments:

COROLLARY 3:* Either the feedback process is not a Poisson process or the arrival process and feedback process are not independent processes (or both) for the *M/M/1* queue with instantaneous, Bernoulli feedback.

PROOF: This result follows immediately from Burke's result [2] on the distribution of the interinput arrivals. For if the feedback process is both independent of the arrival process and is itself a Poisson process, the input process is Poisson. Thus, Burke's result contradicts the assumption. □

3.3 Feedback

The feedback stream seems to be quite difficult to work with. From the previous section we know that it is either not independent of the arrival stream or not a Poisson stream. Melamed [9] has shown that this feedback process is not a Poisson process. We conjecture further that it is not independent of the arrival process. If so, then the known superposition theorems cannot be used to study feedbacks in terms of the arrival, feedback and input processes.

Since the feedback stream is the result of applying a filter to the Markov-renewal output process, it is itself Markov-renewal on the state space $\{1, 2, \ldots\}$. Even in the $M/M/1$ case, the form of the feedback stream does not appear to reduce to that of any simpler process.

4. CONCLUSIONS

There are several conjectures that one can pose concerning networks based on the results of this paper. First with respect to queue length, busy period, and departure processes, if one adopts the "outsiders" view [3] these processes appear to be those generated by an $M/G/1$ queue without feedback. However, if one adopts the "insider" view the queue length process does not appear to behave as seen by the "outsider."

Flow processes in this network cannot be explained by appeal to superposition, stretching, and thinning results for Poisson processes. The requisite independence assumptions both within and between streams of events are not satisfied here. Thus, one cannot assume that these queues which act "as if" they were $M/M/1$ queues to the "outsider" are $M/M/1$ queues to the "insider." In particular, this hints at the possibility that in these networks, even as simple as Jackson networks, any attempt to decompose the network into independent $M/M/1$ queues is doomed to failure. This decomposition must account for the internal flows and these not only appear to be non Poisson, they are nonrenewal and are dependent on each other.

*Melamed [9] has shown, using other arguments, that the feedback stream is not a renewal process.

In [9], it is shown that in the Jackson structure, the flow along any path that returns a customer to a node that he has previously visited is not only not Poisson, it is not renewal. Thus, if Jackson networks have loops, (direct feedback as in this paper being the simplest example), they cannot be decomposed into sub-networks of simple $M/M/1$ queues. In particular, these results imply that a node-by-node analysis of waiting times depending as they do on the "insiders" view is not valid if one simply uses $M/M/1$ results at each server. Takács [10] studies the waiting time problems in the system discussed in this paper. Disney [6] presents another view of the same problem.

ACKNOWLEDGMENTS

We would like to thank Dr. Robert Foley and Dr. Robert B. Cooper for their helpful comments on this paper. In particular Foley first brought to our attention that $\{Q_3^*(n)\}$ and $\{Q_4^*(n)\}$ are asymptotically, identically distributed if one does not include the feedback customers in the queue length. This point was made in his paper [7].

REFERENCES

[1] Buck, R.C., *Advanced Calculus*, (McGraw-Hill, New York, 1956).

[2] Burke, P.J., "Proof of a Conjecture on the Interarrival-Time Distribution in an *M/M/1* Queue with Feedback," IEEE Transactions on Communications, 575-576 (May 1976).

[3] Cooper, R.B., *Introduction to Queueing Theory*, (MacMillan, New York, 1972).

[4] d'Avignon, G.R., and R.L. Disney, "Queues with Instantaneous Feedback," Management Science, *24*, 168-180 (1977).

[5] Disney, R.L., R.L. Farrell, and P.R. deMorais, "Characterization of *M/G/1* Queues with Renewal Departure Processes," Management Science, *19*, 1222-1228 (1973).

[6] Disney, R.L., "Sojourn Times in *M/G/1* Queues with Instantaneous, Bernoulli Feedback," *Proceedings Conference on Point Processes and Queueing Theory*, Keszthely, Hungary (September 1978).

[7] Foley, R.D., "On the Output Process of an *M/M/1* Queue with Feedback," Talk given at San Francisco Meeting, Operations Research Society of America (May 1977).

[8] Jackson, J.R., "Networks of Waiting Lines," Operations Research, *5*, 518-521 (1957).

[9] Melamed, B., "Characterizations of Poisson Traffic Streams in Jackson Queueing Networks," Advances in Applied Probability, *11*, 422-438 (1979).

[10] Takács, L., "A Single Server Queue with Feedback," Bell System Technical Journal, *42*, 509-519 (1963).

AN INVENTORY MODEL WITH SEARCH FOR BEST ORDERING PRICE*

Kamal Golabi

Woodward-Clyde Consultants
San Francisco, California

ABSTRACT

This paper presents a single-item inventory model with deterministic demand where the buyer is allowed to search for the most favorable price before deciding on the order quantity. In the beginning of each period, a sequential random sample can be taken from a known distribution and there is a fixed cost per search. The decision maker is faced with the task of deciding when to initiate and when to stop the search process, as well as determining the optimal order quantity once the search process is terminated. The objective is to minimize total expected costs while satisfying all demands on time. We demonstrate that a set of critical numbers determine the optimal stopping and ordering strategies. We present recursive expressions yielding the critical numbers, as well as the minimal expected cost from the beginning of every period to the end of the horizon.

1. INTRODUCTION

This research is an attempt to marry some aspects of search theory and optimal stopping with inventory theory. Following the pioneering work of Stigler [11], [12], searching for the lowest price is considered a basic feature of economic markets. By citing examples based on real data, Stigler [11] asserted that prices change with varying frequency in all markets, and unless a market is completely centralized, the buyer will not know for certain the prices that the various sellers quote at any given time. This suggests that at any time there will be a frequency distribution of the prices quoted by sellers. If the dispersion of price quotations by sellers is large compared to the cost of search, it will pay—on average—to obtain price quotations from several sellers before taking an "action." The vast literature on search theory (a survey of which can be found in Lippman and McCall [8], DeGroot [5], and Rothschild [10]) is concerned with rules that the searchers should follow when the "action" is accepting or rejecting a price. Once the price has been accepted, the decision process terminates. In many dynamic models, the action is more complicated. In inventory models, for example, the decision not only involves accepting or rejecting an ordering price but how much to order, an action which will affect the search and ordering policies in future periods. In this paper we study such a model. We seek the best search and ordering policies for a simple dynamic inventory problem with deterministic demands where, in the beginning of each period, the purchaser can search for the lowest price before placing an order.

*This research was partially supported by the National Science Foundation through Grant NSF ENG74-13494 and the Air Force Office of Scientific Research (AFOSR 72-2349C).

Classical optimal search considers the following problem: A purchaser can take a sequential random sample $X_1, X_2, \ldots$ from a continuous distribution with a known distribution function F. There is a fixed cost s per observation. Suppose that if the decision maker stops the sampling (search) process after the values $X_1 = x_1, X_2 = x_2, \ldots, X_n = x_n$ have been observed, his cost is $x_n + sn$. Hence, the problem is to find a stopping rule which minimizes $E(X_N + sN)$ where N indicates the random number of observations that are taken under a specified stopping rule. It can be shown that, whether sampling is with or without recall, the optimal stopping rule is characterized by a unique critical number v^* (usually called the reservation price) so that an optimal sampling rule is to continue sampling whenever an observed value exceeds v^* and to stop the process as soon as some observed value does not exceed v^*. Various versions of this problem have been studied by MacQueen and Miller [9], Derman and Sacks [6] and Chow and Robbins [2], [3] among others.

The above search model can be visualized as a one period purchasing problem in which one unit of some commodity has to be purchased at the beginning of the period. Now consider a dynamic multiperiod version of this problem where a demand of one unit has to be satisfied in *each* period and inventory holding cost is charged for items held over for use in subsequent periods. As in the classical search problem, in the beginning of each period a sequential random sample $X_1, X_2, \ldots$ can be taken from a distribution with known distribution function F, but the decision process is not terminated as soon as an acceptable value is observed. The decision maker is faced with the task of deciding how much to order so as to minimize total expected costs while satisfying all demands on time. When the inventory level is sufficient to satisfy the immediate demand, he has also the burden of deciding whether to initiate search at all. This multiperiod model is the subject of our study in this paper.

In Section 2, we present the model. In Section 3, we give the optimal search policy and in Section 4, the optimal ordering policy. We show the intuitive result that an optimal strategy prescribes that search should be initiated only when the inventory level is zero. Furthermore, we show that the reservation price property of the classical search problem still holds. That is, when the inventory level is zero (and therefore search has to be initiated) and n periods remain to the end of the problem, there exists a reservation price α_n such that a price should be accepted if it does not exceed α_n and rejected otherwise. In Section 4, we show that once a price has been accepted, a finite number of critical numbers specify the optimal strategy. The critical numbers divide the interval $[0, \alpha_n]$ into segments so that the interval in which the accepted price falls determines the optimal order quantity. We give recursive expressions which yield α_n as well as the minimal expected cost for any period to the end of the horizon. We will also obtain expressions describing the critical numbers when the holding cost function is convex.

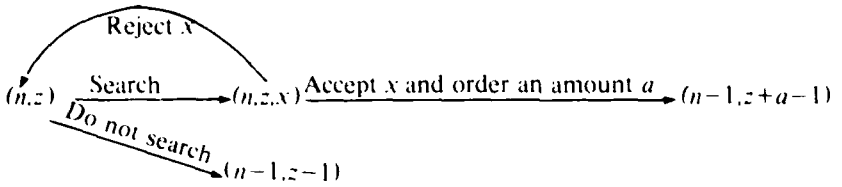
2. THE MODEL

Consider a *multi-period single-item* inventory model in which a demand of one unit has to be satisfied in the beginning of each period and an inventory holding cost is charged. In each period, a sequential random sample $Y_1, Y_2, \ldots$ of ordering prices can be generated from a continuous distribution with known cumulative distribution function $I(\cdot)$, $I(Y_i) < \infty$, and the Y_i's are mutually independent. The cost of generating each random price is s and there is no *limit on* the number of observations which can be made in each period. After receiving a price, the decision maker has to decide whether to accept that price or generate another offer. If he accepts the offered price, he is faced with the decision of how much to order. When the inventory level is sufficient to satisfy the immediate demand, he also has to decide whether to initiate search at all. The objective is to minimize the total expected costs.

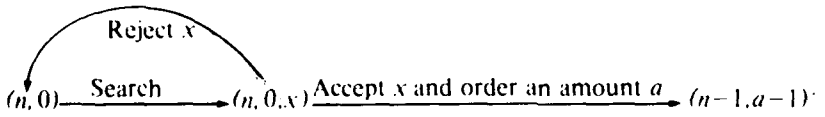
We assume that the length N of the planning horizon is finite, initial inventory is zero, backlogging of demand is not allowed, the cost of holding z units for one period, $h(z)$, is non-decreasing in z and $h(0) = 0$, the purchasing cost is linear in the quantity ordered, and only integer quantities can be ordered. We also assume prices that are not accepted immediately are lost; in view of our results, sampling with recall (of prices in the same period) extends no additional advantage over sampling without recall, and hence would not affect the search policy. Note that when $N = 1$, this model reduces to the classical search problem.

Let n be the number of periods remaining until the end of the horizon, z the inventory on hand with n periods remaining and x the last price received. In each period, our state space consists of numbers (z) and pairs (z,x) corresponding respectively to the state of the system before a search is placed and the state when a search has been placed and an offer x has been received. A policy for period n prescribes a search decision for state (z), and a reject-accept and ordering decision for state (z,x). We assume that for each period an optimal policy exists. Moreover, we restrict our attention to history-independent policies; that is, once the price x has been rejected, we are in the same position as having not placed the search at all. Schematically, remembering that demand in each period equals one, the period-state pairs correspond to each other as follows:

For $z \geqslant 1$:



and



3. OPTIMAL SEARCH POLICY

In this section, we present the optimal search policy. We show that search should only be initiated when the inventory level is zero, and prove that in each period a single reservation price determines the stopping rule. We also give a recursive expression which describes the sequence of reservation prices.

To being, define

$V_n(z,x)$ = the minimal (conditional) expected cost during the last n periods when the inventory level with n periods remaining is z and the last price offered is x

$v_n(z)$ = the minimal expected cost during the last n periods before the decision to search for an offer is made, and when the inventory level with n periods remaining is z.

$u_n(z)$ = the minimal expected cost during the last n periods after the decision to search for an offer in this period has been made, and when the inventory level with n periods remaining is z.

$w_n(z)$ = the minimal expected cost during the last n periods after the decision not to search for an offer in this period has been made, and when the inventory level with n periods remaining is z, $z \geqslant 1$.

$H(z)$ = the total holding cost of z units to be used in z consecutive periods.

Hence, we will have the following relationships:

(1) $$v_n(z) = \min[u_n(z),\ w_n(z)]$$

(2) $$V_n(z,x) = \min\{v_n(z),\ \min_{a \in \{1,2,\ldots,n-z\}} [ax + h(z + a - 1) + v_{n-1}(z + a - 1)]\},$$

(3) $$u_n(z) = s + E_x[V_n(z,x)],$$

(4) $$w_n(z) = h(z - 1) + v_{n-1}(z - 1),\ z \geqslant 1,$$

and

(5) $$H(z) = \sum_{i=1}^{z-1} h(z - i) = \sum_{i=1}^{z-1} h(i).$$

Define

(6a) $$I_n[x,a] \equiv ax + h(a - 1) + v_{n-1}(a - 1)$$

and

(6b) $$I_n(x) \equiv \min_{a \in \{1,2,\ldots,n\}} I_n(x,a),$$

and let $a_n(x)$ be the minimizing value of a in (6a), that is,

(6c) $$I_n(x) = I_n[x,a_n(x)].$$

The quantity $I_n(x)$ is the minimal expected cost attainable during the last n periods when the inventory level with n periods remaining is zero and it has been decided to accept x, the last price offered.

At this point it is natural to ask whether when n periods remain, there exists a single critical price α_n which dictates the acceptance or rejection of a price x when the inventory level is zero. In other words, is there an α_n such that it is optimal to accept the price x (and order a positive amount) if $x \leqslant \alpha_n$ and to continue the search if $x > \alpha_n$. That this is indeed the case, is verified in Theorem 1.

Define

(7) $$\alpha_n \equiv I_n^{-1}[v_n(0)],$$

and the sequence $\{A_n\}_{n=0}^{\infty}$ by the following recursion:

(8a) $$A_0 = 0$$

and

(8b) $$A_n F(\alpha_n) = s + \int_0^{\alpha_n} \min_{a \in \{1,2,\ldots,n\}} [ax + H(a) + A_{n-a}]dF(x) \quad \text{for } n \geqslant 1.$$

We will show later that α_n exists and that A_n equals $v_n(0)$, so that $\alpha_n = I_n^{-1}(A_n)$. These properties are exploited to verify that an optimal policy prescribes that search be initiated, and orders be placed, only when the inventory level is zero. Furthermore, we will show that if the set of prices at which it is optimal to order one unit is nonempty, $\alpha_n = A_n - A_{n-1}$ so that Equation (8b) can be written as

$$A_n F(A_n - A_{n-1}) = s + \int_0^{A_n - A_{n-1}} \min_{a \in \{1,2,\ldots,n\}} [ax + H(a) + A_{n-a}]dF(x), \tag{9}$$

enabling us to obtain the minimal expected cost from the beginning of any period to the end of the horizon by finding $\{A_i\}_{i=0}^{N}$, the unique set of solutions to Equation (9).

THEOREM 1: If the inventory level with n periods remaining is zero, it is optimal to continue the search if x, the last price offered, is greater than α_n and accept the price if $x \leqslant \alpha_n$, where $n = 1, 2, \ldots, N$.

PROOF: Clearly, $I_n(x,a)$ is continuous in x for each n and a, and therefore, $I_n(x)$ is a continuous function of x. Furthermore, for all positive numbers ϵ,

$$I_n(x + \epsilon) = I_n[x + \epsilon, a_n(x + \epsilon)] > I_n[x, a_n(x + \epsilon)] \geqslant I_n[x, a_n(x)] = I_n(x),$$

and hence $I_n(x)$ is strictly increasing in x. Let $\alpha_n(y)$ be such that $I_n[\alpha_n(y)] = y$, i.e., $\alpha_n(y) = I_n^{-1}(y)$. Since

$$v_n(0) \geqslant v_{n-1}(0) \geqslant \min_{a \in \{1,2,\ldots,n\}} [h(a-1) + v_{n-1}(a-1)] = I_n(0),$$

it follows that $\alpha_n = \alpha_n[v_n(0)]$ exists and, as $I_n(x)$ is strictly increasing in x, it is unique (see Figure 1). The first inequality of the above expression follows from the fact that for the $n-1$ period problem we can always follow the optimal policy for the n period problem, so that at each stage m, $n - 1 \geqslant m \geqslant 1$, we would adopt the action prescribed by the n period optimal policy for stage $m + 1$. Thus, the expected cost for the $n-1$ period problem under this policy, $v'_{n-1}(0)$, would be equal to the expected cost of the first $n-1$ periods of the n period problem, and hence $v'_{n-1}(0) \leqslant v_n(0)$. Since $v_{n-1}(0) \leqslant v'_{n-1}(0)$, it follows that $v_n(0)$ is nondecreasing in n.

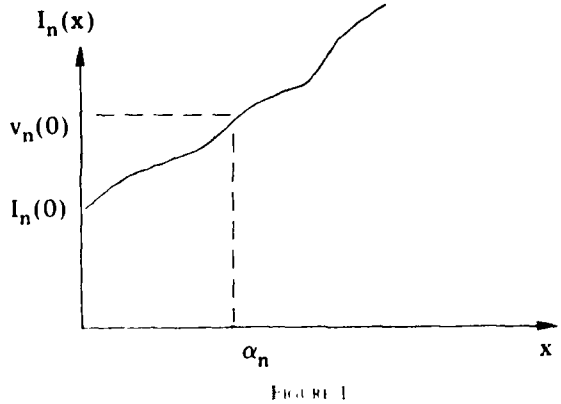


FIGURE 1

From (2) and (6) we have

$$V_n(0,x) = \min [v_n(0), I_n(x)].$$

If $x \leqslant \alpha_n$, then $I_n(x) \leqslant v_n(0)$ so that $V_n(0,x) = I_n(x)$ and search terminates. If $x > \alpha_n$, then $I_n(x) > v_n(0)$ so that $V_n(0,x) = v_n(0)$, in which case it is optimal to continue the search.

Q.E.D.

Thus, when the inventory level is zero, a single critical number determines whether a price should be accepted. We are also interested in finding the optimal strategy when the inventory level is positive. It seems intuitive that if the immediate demand can be satisfied by the current inventory, it would be best to postpone the search—since it is possible to incur the same amount of expected search cost in a later period while saving on the holding cost. The next result, the proof of which is given in the Appendix, verifies this observation. In addition, it shows that the expected cost from any period k in which the inventory level is zero to the end of the horizon equals A_k. Thus, the expected cost from any period can be determined by computing the sequence $\{A_n\}$ from Equation (8b).

THEOREM 2. Under the assumptions of the model, for all k, $k = 1, 2, \ldots, N$,

(a) $v_k(0) = A_k$

(b) $v_k(z) = H(z) + v_{k-1}(0)$ for $1 \leqslant z \leqslant k$.

Theorem 2 verifies that search should be initiated only when the inventory level is zero, and Theorem 1 gives a rule for accepting or rejecting an offered price once search is initiated. These two results however, do not completely specify the optimal strategy. Given that an acceptable price is received, we would like to know how much should be purchased at that price. This question is investigated in the next section.

4. OPTIMAL ORDERING POLICY

In this section we present the optimal ordering policy once an acceptable price has been received. In Corollary 3 we show that a nonincreasing sequence of critical numbers characterize the optimal order quantity. In other words, once a price is received that is less than the reservation price for that period, the interval in which the offered price falls determines the quantity that should be ordered at that price. In Theorem 5 we obtain expressions which describe these critical numbers when the holding cost function is convex.

Before presenting the next result we note that when n periods remain, the inventory level is zero, and an acceptable price has been received, the optimal order quantity is equal to $a_n(x)$. To see this, note that

$$V_n(0,x) = \min[(v_n(0), I_n(x)]$$

by (2) and (6). This fact coupled with Theorem 1 yields $V_n(0,x) = I_n(x)$ whenever $x \leqslant \alpha_n$. Finally, since

$$I_n(x) = I_n[x, a_n(x)] = \min_{1 \leqslant a \leqslant n} [ax + h(a-1) + v_{n-1}(a-1)], \tag{10}$$

it follows that ordering $a_n(x)$ minimizes the expected cost attainable during the last n periods when the inventory level is zero and $x \leqslant \alpha_n$. Note also that by Theorem 2(b), Equation (10) can be written as

$$(11) \qquad I_n(x) = \min_{1 \leqslant a \leqslant n} [ax + H(a) + A_{n-a}].$$

COROLLARY 3: If n periods remain, the inventory level is zero, and an acceptable price has been received, then the optimal order quantity is nonincreasing in the price offered, i.e., $a_n(x') \leqslant a_n(x)$ whenever $x' > x$, $n = 1, 2, \ldots, N$. Consequently, a nonincreasing sequence of critical numbers $\{B_i(n)\}_{i=1}^{n}$ characterize the optimal order quantity. Specifically, it is optimal to order k units whenever $B_k(n) \leqslant x < B_{k-1}(n)$.

PROOF: From (6c) and (11), we have

$$I_n(x) = I_n[x, a_n(x)] = a_n(x) \cdot x + H[a_n(x)] + A_{n-a_n(x)}$$
$$\leqslant I_n[x, a_n(x')] = a_n(x') \cdot x + H[a_n(x')] + A_{n-a_n(x')},$$

giving

$$(12) \qquad x[a_n(x') - a_n(x)] \geqslant A_{n-a_n(x)} - A_{n-a_n(x')} + H[a_n(x)] - H[a_n(x')].$$

If $a_n(x') > a_n(x)$, then (12) implies

$$x'[a_n(x') - a_n(x)] > A_{n-a_n(x)} - A_{n-a_n(x')} + H[a_n(x)] - H[a_n(x')],$$

which yields

$$I_n(x') = a_n(x') \cdot x' + A_{n-a_n(x')} + H[a_n(x')] > a_n(x) \cdot x' + A_{n-a_n(x)}$$
$$+ H[a_n(x)] = I_n[x', a_n(x)],$$

contradicting the fact that $a_n(x')$ is optimal when x' is offered. Q.E.D.

Intuitively, we would expect that when an offered price equals the critical number α_n, we would be indifferent between ordering one unit and not ordering at all. If this were indeed the case, the expected cost when the price is rejected, $v_n(0)$, would be equal to $\alpha_n + v_{n-1}(0)$ yielding $\alpha_n = A_n - A_{n-1}$. This result could then be used to obtain a simple expression for the $B_k(n)$'s when $h(\cdot)$ is convex. As we will show in Lemma 4, the above result holds if the set of prices at which it is optimal to order one unit is nonempty. Unfortunately, as seen from the following example, this is not always the case.

EXAMPLE 1. Let $n = 5$, $s = 5$, $h(z) = 0$ for all z and the price distribution be such that $P(X = 2) = 1 - \epsilon$, and $P(a \leqslant X \leqslant b) = \frac{\epsilon}{4}(b - a)$ for $0 \leqslant a < b \leqslant 4$, where 2 is excluded from all intervals and ϵ is an arbitrary small number. Suppose the offered price in the beginning of the fifth period is 3.

The expected cost of rejecting the offered price is (approximately)

$$5 + 2 \times 5 = 15,$$

as one would pay the search cost of 5 and almost definitely receive the price of 2, at which one would order 5 units. However, the expected cost of ordering i units, $i \leqslant 4$, is (approximately)

$$3i + 5 + 2(5 - i) = 15 + i,$$

while the cost of ordering 5 units is 15. Hence, we would be indifferent between not ordering and ordering at $x = 3$, which implies that $\alpha_5 = 3$.

Since at $x = 3$ we order 5 units, any price above 3 is rejected, and the optimal order quantity $a_n(x)$ is nonincreasing in x, it follows that $[x : a_n(x) = 1]$ is empty.

LEMMA 4: If $[x : a_n(x) = 1]$ is nonempty, then $\alpha_n = A_n - A_{n-1}$.

PROOF: Let $\bar{x}$ be the largest x such that $a_n(x) = 1$. By Theorem 1, α_n is the highest price at which it is optimal to order a positive quantity. Therefore, $\bar{x} \leqslant \alpha_n$. Consequently, we can conclude from Corollary 3 that $a_n(\alpha_n) \leqslant 1$, but $a_n(\alpha_n)$ is positive so that $a_n(\alpha_n) = 1$. From Theorem 2 and Equations (7) and (11), we have

$$A_n = v_n(0) = I_n(\alpha_n) = \min_{1 \leqslant a \leqslant n} [a\alpha_n + H(a) + A_{n-a}]$$

$$= [a_n(\alpha_n) \cdot \alpha_n + H[a_n(\alpha_n)] + A_{n-a_n(\alpha_n)}] = \alpha_n + H(1) + A_{n-1},$$

which yields $\alpha_n = A_n - A_{n-1}$. Q.E.D.

Whereas we cannot determine in advance the conditions under which Lemma 4 would hold, we can proceed by assuming that the lemma holds, and determine the sequence $\{\bar{A}_n\}_{n=0}^{N}$ that satisfies Equation (9). We then can obtain $\{\bar{\alpha}_n\}$ from $\bar{\alpha}_n = I_n^{-1}(\bar{A}_n)$. If $\{\bar{\alpha}_n\}$ and $\{\bar{A}_n\}$ also satisfy Equation (8b), by uniqueness of the solution, α_n is indeed equal to $A_n - A_{n-1}$.

It is interesting to note that contrary to what one might expect, α_n is not monotone in n. Before Theorem 5, we give examples where α_n is not monotone irrespective of whether $[x : a_n(x) = 1]$ is empty or not.

EXAMPLE 2: (a) Consider again Example 1. Since we would almost definitely receive the price of 2 after the first search, we have

$$v_n(0) = s + \min_{1 \leqslant a \leqslant n} [ax + H(a) + v_{n-a}(0)].$$

Thus,

$$v_1(0) = 5 + 2 = 7$$

$$v_2(0) = 5 + \min (2 + 7, 4) = 9.$$

From $v_n(0) = I_n(\alpha_n)$, we have $\alpha_1 = v_1(0) = 7$ and

$$9 = \min (\alpha_2 + 7, 2\alpha_2)$$

yielding $\alpha_2 = 4.5$. As shown earlier, $\alpha_5 = 3$. Therefore, α_n is not monotone in n.

(b) We note that α_n is not necessarily monotone even if $[x : a_n(x) = 1]$ is nonempty. Consider the case where the price distribution is the same as Example 1. However, there is a holding cost of 1 per unit per period and $s = 2$. Then, $H(1) = 0$, $H(2) = 1$, $H(3) = 3$ and $H(4) = 6$ and

$$v_1(0) = 2 + 2 = 4$$

$$v_2(0) = 2 + \min [4 + 1, 2 + 4] = 7$$

$$v_3(0) = 2 + \min [6 + 3, 4 + 1 + 4, 2 + 7] = 11$$

$$v_4(0) = 2 + \min [8 + 6, 6 + 3 + 4, 4 + 1 + 7, 2 + 11] = 14.$$

From $v_n(0) = I_n(\alpha_n)$, we have

$$4 = \alpha_1$$
$$7 = \min [2\alpha_2 + 1, \alpha_2 + 4]$$
$$11 = \min [3\alpha_3 + 3, 2\alpha_3 + 1 + 4, \alpha_3 + 7]$$
$$14 = \min [4\alpha_4 + 6, 3\alpha_4 + 3 + 4, 2\alpha_4 + 1 + 7, \alpha_4 + 11]$$

yielding

$$\alpha_1 = 4,\ \alpha_2 = 3,\ \alpha_3 = 4,\ \alpha_4 = 3.$$

Note that in this example, $a_n(\alpha_n) = 1$ for $1 \leqslant n \leqslant 4$ and the condition for Lemma 4 holds. It can be easily verified that $\alpha_n = v_n(0) - v_{n-1}(0)$ for all $1 \leqslant n \leqslant 4$.

THEOREM 5: If the condition for Lemma 4 holds and if the holding cost function $h(\cdot)$ is convex, then

$$B_k(n) = \alpha_{n-k} - h(k), \quad \text{where } 1 \leqslant k \leqslant n. \tag{13}$$

PROOF: We have to show that

(a) The RHS of (13) is nonincreasing in k.

(b) It is optimal to order k units if x, the price offered, satisfies

$$\alpha_{n-k} - h(k) \leqslant x \leqslant \alpha_{n-(k-1)} - h(k-1). \tag{14}$$

To show (a), we note that

$$A_{n-k+1} = v_{n-k+1}(0) = I_{n-k+1}(\alpha_{n-k+1}) = \min_{1 \leqslant a \leqslant n-k+1} [a\alpha_{n-k+1} + H(a) + A_{n-k+1-a}] \leqslant 2\alpha_{n-k+1} + H(2) + A_{n-k-1}$$
$$= 2(A_{n-k+1} - A_{n-k}) + h(1) + A_{n-k-1},$$

where the first equality follows from Theorem 2, the second from (7), the third from (11) and the last from Lemma 4. Thus, by convexity of $h(\cdot)$,

$$h(k) - h(k-1) \geqslant h(1) \geqslant A_{n-k} - A_{n-k-1} - A_{n-k+1} + A_{n-k}$$
$$= \alpha_{n-k} - \alpha_{n-k+1}$$

and, therefore, (a) is true.

To show (b), suppose x is such that (14) holds. We show that $I_n(x,k-j) \leqslant I_n(x,k-j-1)$ for each $j \geqslant 0$, and therefore ordering k units is at least as good as ordering any amount less than k. Suppose $I_n(x,k-j) > I_n(x,k-j-1)$. Then

$$(k-j-1)x + A_{n-(k-j-1)} + H(k-j-1) < (k-j)x + A_{n-(k-j)} + H(k-j)$$

which yields

$$x > A_{n-(k-j-1)} - A_{n-(k-j)} - h(k-j-1)$$
$$= \alpha_{n-(k-j-1)} - h(k-j-1) \geqslant \alpha_{n-(k-1)} - h(k-1),$$

where the last inequality follows from (a). This contradicts the right inequality of (14). Therefore, $I_n(x,k-j) \leqslant I_n(x,k-j-1)$.

$L_n(x,k+j) \leq L_n(x,k+j+1)$ for each $j \geq 0$ by a similar proof.

Hence, it is optimal to order k units whenever (14) holds. Q.E.D.

5. REMARKS

The purpose of this study has been to investigate optimal search policies in the context of a sequential model. The underlying inventory model has been chosen as a rather simple one. There are no setup costs involved and the demand equals one unit in each period. It would be interesting to investigate more general problems. We suspect that both the reservation price property of Theorem 1 and the Wagner-Whitin [13] type result of Theorem 2 (order only when current inventory level is zero) would still hold for models with setup costs and arbitrary deterministic demands. The optimal policy would be a function of setup costs as well as the holding cost and price distribution. The results should also hold when the price distributions are nonstationary. Given that the initial inventory is zero, the ordering policy will be such that there is no inventory in the beginning of periods with favorable price distributions.

Another interesting extension is the case wherein the search process is adaptive. The searcher does not know the exact distribution of price; the price offer is used not only as an opportunity to order at that price but also as a piece of information to update the prior distribution. When the distribution of prices is not known exactly, the form of the optimal policy is not obvious. As Rothschild [10] points out, the reservation price property of Theorem 1 would not necessarily hold even for a one period problem. Rothschild presents the following example. Suppose there are three prices, \$1, \$2, and \$3, and that the cost of search is \$0.01. Prior beliefs admit the possibility of only two distributions of prices. Either all prices are \$3 or they are distributed between \$1 and \$2 in the proportions 99 to 1. A man with these beliefs should accept a price of \$3 (as this is a signal that no lower prices are to be had) and reject a quote of \$2 (which indicates that the likelihood that a much better price will be observed on another draw is high).

However, when the distribution is a member of certain families of distributions but has one or more unknown parameters, Rothschild [10], DeGroot [5] and Albright [1] have shown that the reservation price property holds for the one-period problem. We conjecture that when the distribution of price is stationary but is not known exactly, search should be initiated only when the inventory level is zero. If this is the case and the distribution belongs to one of the families of distributions studied by Rothschild [10] and Albright [1], then the reservation price property as well as the ordering policy presented in Section 4 should still hold.

ACKNOWLEDGMENTS

This paper is essentially Chapter 3 of the author's dissertation (1976) at the University of California, Los Angeles. The author expresses his appreciation to Professor Steven Lippman for his guidance and encouragement. He also appreciates several helpful comments by Professor Sheldon Ross and the referee.

BIBLIOGRAPHY

[1] Albright, C.S., "A Bayesian Approach to a Generalized House Selling Problem," Management Science *24*, 432-440 (1977).

[2] Chow, Y.S. and H. Robbins, "A Martingale System Theorem and Applications," *Proceedings of the 4th Berkeley Symposium on Mathematical Statistics and Probability*, University of California Press, Berkeley, California (1961).

[3] Chow, Y.S. and H. Robbins, "On Values Associated with a Stochastic Sequence," *Proceedings of the 5th Berkeley Symposium on Mathematical Statistics and Probability*, University of California Press, Berkeley, California (1967).
[4] DeGroot, M.H., "Some Problems of Optimal Stopping," Journal of the Royal Statistical Society *30*, 108-122 (1968).
[5] DeGroot, M.H., *Optimal Statistical Decisions* (McGraw-Hill, 1970).
[6] Derman, C. and J. Sacks, "Replacement of Periodically Inspected Equipment," Naval Research Logistics Quarterly *7*, 597-607 (1960).
[7] Golabi, K., "Optimal Inventory and Search Policies with Random Ordering Costs," Working Paper No. 252, Western Management Science Institute, University of California, Los Angeles (1976).
[8] Lippman, S.A. and J.J. McCall, "The Economics of Job Search: A Survey," Economic Inquiry *14*, 155-189 (1976).
[9] MacQueen, J.B. and R.G. Miller, Jr., "Optimal Persistence Policies," Operations Research *8*, 362-380 (1960).
[10] Rothschild, M., "Models of Market Organization with Imperfect Information: A Survey," Journal of Political Economy *81*, 1283-1308 (1973).
[11] Stigler, G.J., "The Economics of Information," Journal of Political Economy *69*, 213-225 (1961).
[12] Stigler, G.J., "Information in the Labor Market," Journal of Political Economy *70*, 94-104 (1962).
[13] Wagner, H.M. and T.M. Whitin, "Dynamic Version of the Economic Lot Size Model," Management Science *5*, 89-96 (1958).

APPENDIX

THEOREM 2: Under the assumptions of the model, for all k, $k = 1, 2, \ldots, N$,

(a) $v_k(0) = A_k$

(b) $v_k(z) = H(z) + v_{k-z}(0)$ for $1 \leqslant z \leqslant k$.

Consequently, the search process is initiated only when the inventory level is zero.

Before proving Theorem 2, we establish two elementary facts.

LEMMA A: For any two positive integers i and j, $H(i + j) \geqslant H(i) + H(j)$.

PROOF:

$$H(i+j) = \sum_{k=1}^{i+j-1} h(k) = \sum_{k=1}^{i-1} h(k) + \sum_{k=i}^{i+j-1} h(k) \geqslant \sum_{k=1}^{i-1} h(k) + \sum_{k=1}^{j-1} h(k)$$
$$= H(i) + H(j). \qquad \text{Q.E.D}$$

LEMMA B: The integral $\int_0^{\alpha_n(y)} [y - I_n(x)]\, dF(x) \equiv G_n(y)$ is strictly increasing in y, continuous, and unbounded above.

PROOF: Since $I_n[\alpha_n(y)] = y$ and $I_n(x)$ is strictly increasing in x, it follows that $\alpha_n(y)$ is strictly increasing in y. Hence, $G_n(y)$ is strictly increasing, continuous (as F is continuous) and unbounded above.

Q.E.D.

PROOF OF THEOREM 2: The proof is by induction on k. From Equations (1), (3), (2) and (6), we have

$$\text{(A-1)} \qquad \begin{aligned} v_k(0) = u_k(0) &= s + E_x[V_k(0,x)] \\ &= s + E \min \left\{ v_k(0), \min_{1 \le a \le k} [ax + h(a-1) + v_{k-1}(a-1)] \right\} \\ &= s + E \min [v_k(0), I_k(x)]. \end{aligned}$$

For $k = 1$, (b) is obvious. To show (a), note that by (6), $I_1(x) = x$. Next, from (A-1) we have

$$v_1(0) = s + E \min [v_1(0), x] = s + \int_0^{v_1(0)} x dF(x) + \int_{v_1(0)}^{\infty} v_1(0)\, dF(x),$$

from which we obtain

$$\text{(A-2)} \qquad v_1(0)\, F[v_1(0)] = s + \int_0^{v_1(0)} x dF(x).$$

(Note the close connection between $v_1(0)$ and the maximizing price in the house selling problem.) In order to determine whether $v_1(0)$ is the unique solution to (A-2), note that it is equivalent to verify that $s = \int_0^y (y - x)\, dF(x) = G_1(y)$ has a unique solution. The latter result follows from Lemma B.

From (7) we have $I_1(\alpha_1) = v_1(0)$ and therefore $\alpha_1 = v_1(0)$. Thus, (A-2) becomes

$$\alpha_1 F(\alpha_1) = s + \int_0^{\alpha_1} x dF(x),$$

which coupled with (8b) for $n = 1$, gives $A_1 = \alpha_1 = v_1(0)$ so that (a) holds for $k = 1$.

Assume (a) and (b) hold for $k = 1, 2, \ldots, n-1$. We show that the theorem holds for $k = n$.

From (A-1), we have

$$\begin{aligned} v_n(0) &= s + E \min [v_n(0), I_n(x)] \\ &= s + \int_0^{\alpha_n} I_n(x)\, dF(x) + \int_{\alpha_n}^{\infty} v_n(0) dF(x) \\ &= [F(\alpha_n)]^{-1} \left\{ s + \int_0^{\alpha_n} \left(\min_{1 \le a \le n} [ax + h(a-1) + v_{n-1}(a-1)] \right) dF(x) \right\} \\ &= [F(\alpha_n)]^{-1} \left\{ s + \int_0^{\alpha_n} \left(\min_{1 \le a < n} [ax + h(a-1) + H(a-1) + v_{n-a}(0)] \right) dF(x) \right\} \\ &= [F(\alpha_n)]^{-1} \left\{ s + \int_0^{\alpha_n} \left(\min_{1 \le a \le n} [ax + H(a) + A_{n-a}] \right) dF(x) \right\} \\ &= [F(\alpha_n)]^{-1} A_n F(\alpha_n) \\ &= A_n, \end{aligned}$$

where the second equality follows from Theorem 1, the third from a simple rearrangement of terms, the fourth and fifth equalities from the induction hypothesis and the sixth equality from (8b). Therefore (a) is true for $k = n$.

Since we are assuming that (b) holds for $k = n - 1$, it follows that

$$\text{(A-3)} \qquad I_n(x) = \min_{1 \le a \le n} [ax + H(a) + A_{n-a}]$$

and

$$A_n F(\alpha_n) = s + \int_0^{\alpha_n} I_n(x)\, dF(x),$$

which gives

$$\text{(A-4)} \qquad \int_0^{\alpha_n} [A_n - I_n(x)]\, dF(x) = s.$$

We note that by (4) and the induction hypothesis, for $1 \leqslant z \leqslant n$

$$\begin{aligned} w_n(z) &= h(z-1) + v_{n-1}(z-1) = h(z-1) + H(z-1) + v_{n-z}(0) \\ &= H(z) + v_{n-z}(0), \end{aligned}$$

and therefore to prove (b) for $k = n$, it suffices to show $v_n(z) = w_n(z)$ whenever $z \geqslant 1$. That is, we need to show $u_n(z) \geqslant H(z) + v_{n-z}(0)$ whenever $z \geqslant 1$.

We can write

$$\begin{aligned} u_n(z) &= s + E \min\left\{v_n(z), \min_{1 \le a \le n-z} [ax + h(z+a-1) + v_{n-1}(z+a-1)]\right\} \\ &= s + E \min\left\{u_n(z), w_n(z), \min_{1 \le a \le n-z} [ax + h(z+a-1) + v_{n-1}(z+a-1)]\right\} \\ &= s + E \min\left\{u_n(z), \min_{0 \le a \le n-z} [ax + h(z+a-1) + v_{n-1}(z+a-1)]\right\} \\ &= s + E \min\left\{u_n(z), \min_{0 \le a \le n-z} [ax + H(z+a) + A_{n-z-a}]\right\} \\ &\geqslant s + E \min\left\{u_n(z), H(z) + \min_{0 \le a \le n-z} [ax + H(a) + A_{n-z-a}]\right\} \\ &= s + E \min\left\{u_n(z), H(z) + \min [A_{n-z}, \min_{1 \le a \le n-z} (ax + H(a) + A_{n-z-a})]\right\} \\ &= s + E \min\left\{u_n(z), H(z) + \min [v_{n-z}(0), I_{n-z}(x)]\right\}, \end{aligned}$$

where the first equality follows from (3) and (2), the second from (1), the fourth from induction hypothesis and the last equality from the induction hypothesis and (A-3). The inequality follows from Lemma A. Hence,

$$\text{(A-5)} \qquad \gamma \equiv u_n(z) - H(z) \geqslant s + E \min\left\{u_n(z) - H(z), \min [v_{n-z}(0), I_{n-z}(x)]\right\}.$$

If γ were less than $v_{n-z}(0)$, then from (A-5) we would have

$$\gamma \geqslant s + E \min [\gamma, I_{n-z}(x)]$$

giving

$$\int_0^{\alpha_{n-2}(\gamma)} [\gamma - l_{n-2}(x)]dF(x) \geqslant s = \int_0^{\alpha_{n-2}} [A_{n-2} - l_{n-2}(x)]dF(x),$$

where the equality follows from (A-4). Hence,

$$G_{n-2}(\gamma) \geqslant G_{n-2}(A_{n-2}) = G_{n-2}[v_{n-2}(0)],$$

contradicting Lemma B. Therefore, $\gamma \geqslant v_{n-2}(0)$, which completes the induction argument.

Q.E.D.

THE UNITED STATES COAST GUARD COMPUTER-ASSISTED SEARCH PLANNING SYSTEM (CASP)*

Henry R. Richardson

Daniel H. Wagner, Associates
Paoli, Pennsylvania

Joseph H. Discenza**

U.S. Coast Guard

ABSTRACT

This paper provides an overview of the Computer-Assisted Search Planning (CASP) system developed for the United States Coast Guard. The CASP information processing methodology is based upon Monte Carlo simulation to obtain an initial probability distribution for target location and to update this distribution to account for drift due to currents and winds. A multiple scenario approach is employed to generate the initial probability distribution. Bayesian updating is used to reflect negative information obtained from unsuccessful search. The principal output of the CASP system is a sequence of probability "maps" which display the current target location probability distributions throughout the time period of interest. CASP also provides guidance for allocating search effort based upon optimal search theory.

1. INTRODUCTION

This paper provides an overview of the computer-assisted search planning (CASP) system developed for the United States Coast Guard to assist its search and rescue (SAR) operations. The system resides on a CDC 3300 located in Washington, D.C., and can be used by all USCG Rescue Coordination Centers (RCCs) in the continental United States and Hawaii via remote access terminals.

The Coast Guard is engaged daily in search and rescue missions which range from simple to complex. The amount of information available to predict the position of the search target ranges from extremely good to almost no information at all. The process of planning, commanding, and evaluating these searches takes place in Rescue Coordination Centers (RCCs) located throughout the United States in major coastal cities.

The entire planning process begins with the awareness that a distress on the water may exist. This awareness usually results from a telephone call from a friend or relative or from a radio communication from the boat or vessel itself.

*This work was supported in part by USCG Contract DOT-CG-52489-A and ONR Contract No. N0001[illegible]
**The opinions or assertions contained herein are the private views of the [illegible] and are not to be construed as official or reflecting the views of the Commandant or the Coast Guard at large.

Next all available information has to be evaluated to decide whether or not to begin a search, and what level of effort is required given the search begins. At this point a great deal of effort goes into deciding where the distress incident occurred. This might be considered the first phase of planning.

The next phase involves computing where the search target will be when the first search units arrive on scene. Among other things, this requires the prediction of ocean drift and wind velocity and the estimation of uncertainties in these predictions.

The next questions pertain to the effort allocation process—how much effort must be expended and in what areas? Prior to the advent of computer search programs, SAR planners relied upon various rules of thumb as presented in the National Search and Rescue Manual [11]. Simplicity was necessary to facilitate hand computation, but at the same time prevented adequate treatment of the many sources of uncertainty which characterize a SAR incident.

The search phase is the actual deployment of aircraft and vessels, the conduct of preset search patterns, and the report of results back to the RCC.

If the search is unsuccessful for that day, then the results must be reevaluated and a new search planned for the following day.

This process continues until the target is found or until the search is terminated. In brief (and in slightly more technical terms), the planning phases are as follows:

(1) Determine the target location probability distribution at the time of the distress incident.

(2) Update the target location probability distribution to account for target motion prior to the earliest possible arrival of a search unit on-scene.

(3) Determine the optimal allocation of search effort, and estimate the expected amount of search effort required to find the target.

(4) Execute the search.

(5) If the search is unsuccessful, evaluate the results and update the target location probability distribution to account for this negative information.

(6) Repeat the planning procedures in Steps (2) through (5) until the target is found or the search is terminated.

These planning phases are illustrated in the CASP case example given in Section 3.

The first efforts at computerization concentrated on the target location prediction process. Oceanographic models were used to compute drift and to estimate target position. The Monterey Search Planning Program and the Coast Guard's own Search and Rescue Planning System, SARP, represented early computer assisted search efforts. Even today, in cases where the information available makes the planning straightforward, the SARP program does nicely.

In 1970, the Office of Research and Development in Washington funded development of a more comprehensive approach to search planning based in part on lessons learned in the Mediterranean H-bomb search in 1966 (Richardson [5]) and in the Scorpion search in 1968

(Richardson and Stone [6]). In 1972, the CASP system was delivered to the Operations Analysis Branch of Commander Atlantic Area in New York for evaluation, implementation, and training. The system was made operational early in 1974.

CASP is now in use in 11 Coast Guard rescue centers. In addition, CASP has been used at the Air Force Central Rescue Headquarters at Scott AFB, Illinois, to help plan and coordinate search missions for lost airplanes within the continental United States. A modification of the CASP system has also been provided to the Canadians for inland SAR planning.

At the present time, the use of CASP is limited to open ocean searches. Even though these searches represent but a small percentage of the total U.S. Coast Guard search operations, CASP has been credited with saving over a dozen lives.

Section 2 provides a description of the CASP methodology. Section 3 illustrates the use of CASP in an actual SAR incident involving the 1976 sinking of the sailing vessel S/V Spirit in the Pacific, and Section 4 describes CASP training.

2. CASP METHODOLOGY

The CASP information processing methodology is based upon Monte Carlo simulation to obtain an initial probability distribution for target location and to update this distribution to account for drift due to currents and winds. A multiple scenario approach is employed to generate the initial probability distribution. In the sense used here, a scenario is a hypothetical description of the distress incident which provides quantitative inputs for the CASP programs. Bayesian updating is used to reflect negative information obtained from unsuccessful search.

The principal output of the CASP system is a sequence of probability "maps" which display the current target location probability distributions throughout the time period of interest. CASP also provides guidance for allocating search effort based upon optimal search theory.

The CASP system is composed of a number of different programs, each designed for a different information processing function. The program components are MAP, POSITION, AREA, TRACKLINE, COMBINATION, DRIFT, RECTANGLE, PATH, and MULTI; the functions are as follows:

(1) display the probability maps (MAP),

(2) generate an initial distribution of target location at the time of distress (POSITION, AREA, TRACKLINE, and COMBINATION),

(3) update the target location probability distributions for motion subsequent to the time of distress (DRIFT),

(4) update the target location probability distributions for negative search results and compute the cumulative detection probability (RECTANGLE and PATH), and

(5) calculate optimal allocations of search effort (MAP and MULTI).

These programs will be described below following presentation of an overview of the general system design.

CASP System Design

The CASP system design was motivated by a desire to provide a highly realistic probabilistic description for the target's location at the time of the distress incident and for the target's substantial motion. In view of the success achieved in the Mediterranean H-bomb search [12] in 1966, and in the Scorpion search [5] in 1968, it seemed evident that a Bayesian approach would provide a practical method for incorporating information gained from unsuccessful search.

Target motion modeling posed a more difficult problem. Models which were amenable to an "analytic" approach were not flexible enough to give a good representation of the search facts. For example, Gaussian motion processes (or mixtures of Gaussian processes) were unsatisfactory in cases where the search facts required a uniform or annular shaped target location probability density. Markov chains based on transitions among search grid cells were unsatisfactory in cases where one desired to change the grid in the course of an operation. In general, these models tended to force the facts to fit the mathematics to an undesirable extent.

It was also desired to develop a modular system so that additional features and improvements could be made as time went on. In order to gain the confidence of the users, the system had to be simple to understand and require a minimum of unfamiliar inputs. The design which seemed best suited in view of the above considerations is a hybrid approach which uses Monte Carlo to simulate target motion and analytic methods to compute detection probabilities.

A motivation for use of Monte Carlo was the recognition that computation of the posterior target location probability distribution can be viewed as the numerical evaluation of a multivariate integral of high dimensionality. In such cases (i.e., high dimensionality), classical numerical integration techniques perform poorly (see, for example, Shreider [7]) especially when the integrands can have jump discontinuities and are not of a simple analytic form. These problems are typical of CASP applications. Discontinuities occur when the "target" moves into a region where search effort is concentrated, and the joint probability density for target position at several specified times during the search is a very complicated function.

The underlying structure of CASP is a Markov process, with a three-dimensional state space consisting of points (X, Y, Φ). The variables X and Y denote latitude and longitude and Φ denotes search failure probability. For $j = 1, \ldots, J$, the jth Monte Carlo replication (X_j, Y_j, Φ_j) represents the target's current position (time is implicit) together with the cumulative probability of search failure for that particular target replication computed for its entire history. Target motion is assumed to be Markovian and successive increments of search are assumed to be statistically independent. Thus (X_j, Y_j, Φ_j) completely describes the state of the jth target replication at a given moment.

Figure 1 provides a schematic diagram for the operation of the CASP system. All of the programs mentioned will be discussed individually in subsequent subsections. The first step is to construct a file (called the "target-triple file") consisting of samples from the target location probability distribution at the time of the distress incident. This file is stored on computer disc and processed sequentially by various programs.

These initial points $(X_j, Y_j, 1)$ have failure probabilities $\Phi_j = 1$, since no search has yet been carried out. The target positions (X_j, Y_j) are sampled from a probability density function F of the form

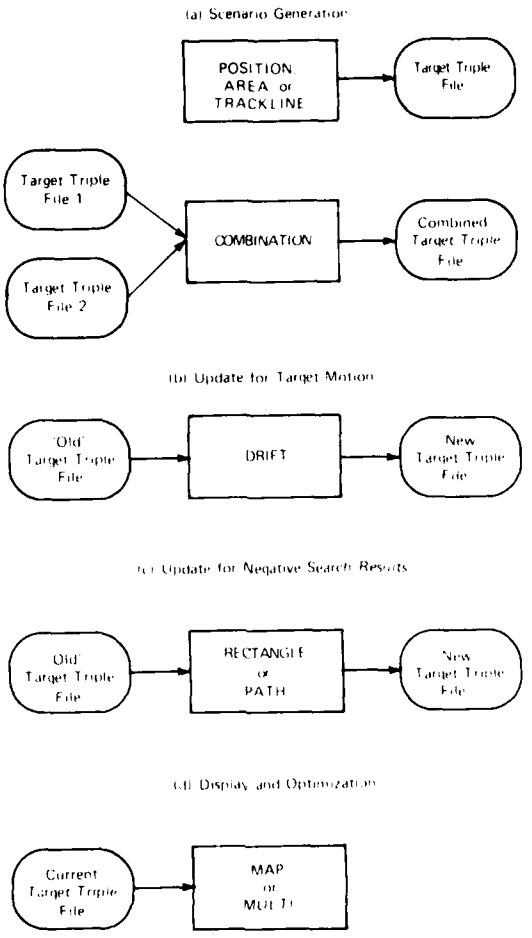


FIGURE 1. Casp system design

$$F = \sum_{k=1}^{K} w_k f_k ,$$

where f_k is the density corresponding to the kth "scenario," and $w_k > 0$ is the scenario's subjectively assigned weight $\left(\sum_{k=1}^{K} w_k = 1\right)$.

Monte Carlo samples from a probability density F are obtained by first using one of the "generation programs" POSITION, AREA, or TRACKLINE. Averages of densities of different types are obtained by forming preliminary target triple files with two or more "generation" programs and then combining them with the program COMBINATION. The construction of the prior target location probability distribution is shown schematically in Figure 1(a).

Updates for target motion (Figure 1(b)) or to account for negative search results (Figure 1(c)) are carried out by reading the "old" target triple file from disc into the appropriate program and outputting a "new" target triple file. When program DRIFT is used (Figure 1(b)), the values of X_i and Y_i are modified, but the value of Φ_i remains unchanged. For an update for

negative search results, the file is first updated for motion by use of program DRIFT. The target triple file is frozen at the mid-search time and then modified by RECTANGLE or PATH. These programs modify Φ_i by use of Bayes' theorem but the position variables X_i and Y_i remain the same since motion is frozen.

The probability distributions and optimal allocations of search effort are displayed using program MAP or MULTI (Figure 1(d)). In both cases, this is a read-only operation, and the target triple file is not modified.

Display

The MAP program displays the target location probability distributions in a two dimensional format. Figure 2 shows an example of a probability map corresponding to an actual SAR case. The geographical region is divided into cells oriented north-south and east-west and the target location probabilities* for each cell are multiplied by 10,000 and displayed. Thus, the number 1800 in a cell indicates that the target location probability is .18. Equal probability contours are usually sketched to make it easier to visualize the probability distribution.

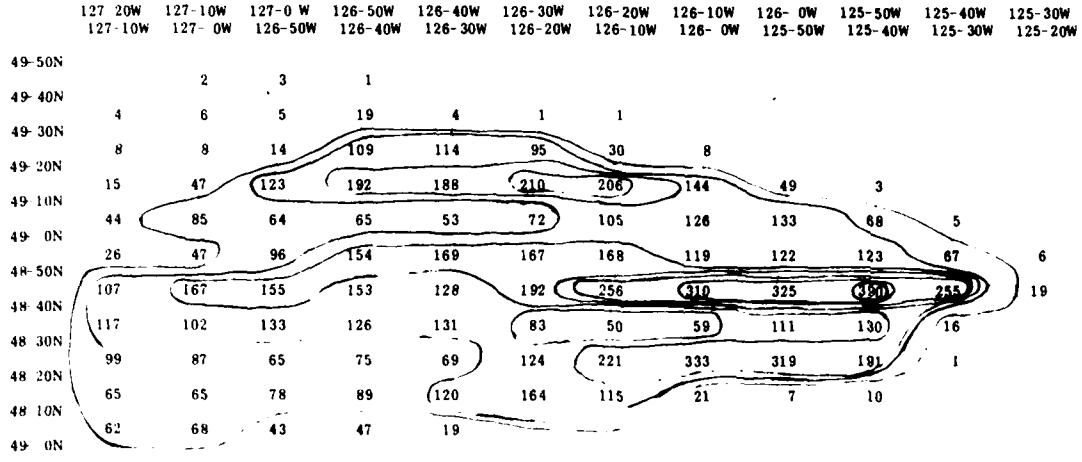


Fig. 2. Target location probability distribution probabilities are multiplied at 10,000 and truncated

A "quick map" in which symbols are used to represent ranges of probabilities can also be output. The quick map provides a compact version of the probability distribution which is suitable for a quick appraisal of the search situation and is convenient for inclusion in after-action reports.

Finally, MAP can output an ordered list of the highest probability cells and the amount of effort to be placed in each cell in order to maximize detection probability. More will be said about search optimization in the last subsection.

*The format implies higher accuracy than is warranted in view of the Monte Carlo procedures employed.

Initial Target Location Probability Distribution

The initial target location probability distribution is constructed from "building block distributions" using a weighted scenario approach. The individual building block distributions are generated by the use of one or more of the programs POSITION, TRACKLINE, and AREA. Program COMBINE is used to combine the outputs of the individual "generation" programs.

In most SAR cases, there is scant information available about the target's position at the time of distress. Sometimes, for example, a fisherman simply is reported overdue at the end of a day. He may have been planning to fish in one of several fishing grounds but did not make his precise intentions known.

In other cases, more information is available. For example, it might be known that a vacationer was intending to sail from one marina to another but never arrived at the intended destination. In some cases, it might also be known that there was bad weather along the intended route. This would make some positions along track more likely for a distress than others.

In order to encourage inclusion of diverse possibilities in these scenarios, it is a recommended practice for two or three search planners to work out the details together. The remainder of this subsection will describe the programs POSITION, AREA, and TRACKLINE which are used to simulate the scenarios and generate the initial target location probability distribution.

Position. A POSITION scenario has two parts, an initial position and a subsequent displacement. POSITION can be used to generate a weighted average of as many as ten scenarios.

The initial position probability distribution is modeled as a bivariate normal distribution, and the displacement is modeled as a distribution over an annular sector. In the latter distribution, the angle and distance random variables are assumed to be independent and uniformly distributed between minimum and maximum values input by the user. The displacement distribution is useful, for example, in cases where the initial position corresponds to the last fix on the target and where one can estimate the course and speed of subsequent movement prior to the occurrence of the distress incident.

The displacement option can also be used in cases involving a "bail out" where it can describe the parachute drift. The amount of displacement in this case will depend upon the altitude of the aircraft and the prevailing winds at the time. Since these factors are rarely known precisely, the capability to "randomize" direction and distance is an important feature.

Area. The second generation program is AREA. This program is used to generate an initial target location probability distribution in cases where a general region can be postulated for the location of the distress incident but where a normal distribution simulated by POSITION would be a poor representation of the uncertainty. Each scenario for program AREA determines a uniform probability density within a convex polygon. AREA might be used, for example, when a lost fisherman's usual fishing ground is known from discussions with friends and relatives. As with POSITION, AREA can generate a weighted average of 10 scenarios.

Trackline. The third and last generation program is TRACKLINE. This program is the most complex of the generation programs and is used when target track information is available from a float plan or some other source. TRACKLINE creates a probability distribution about a base track. This track can be constructed from as many as 10 segments, each of which can be a portion of a rhumb line or of a great circle.

The motion of the target about each base track segment is specified by three circular normal probability distributions corresponding to target position at the initial, mid-point, and endpoint of each segment. Each simulated target track is obtained by drawing random numbers for target position from these distributions and then connecting the points with straight lines.

Figure 3 illustrates a typical situation. The target's point of departure and intended destination are assumed known, and a base track is constructed between these points. The base track might be taken from the target's float plan or hypothesized from the target's past habits. In the case illustrated by Figure 3, there are three track segments. The 50% circles of uncertainty are assumed to grow in size to about midway along the track and then diminish. Since the point of departure and intended destination are assumed to be known, the extreme endpoints of the entire track have zero uncertainty.

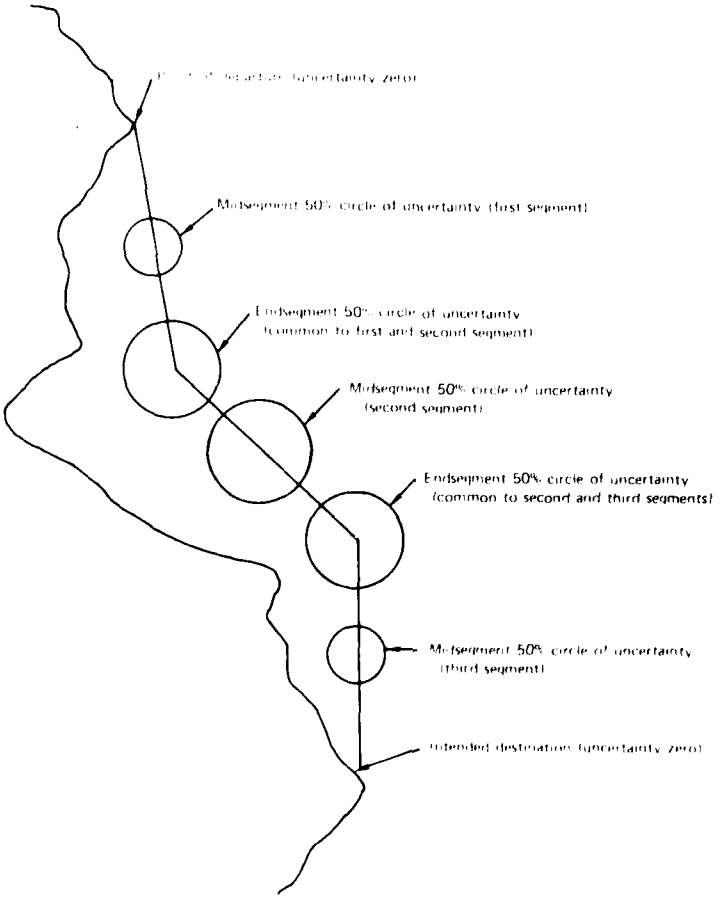


FIGURE 3. Description of trackline uncertainties.

Trackline. The third and last generation program is TRACKLINE. This program is the most complex of the generation programs and is used when target track information is available from a float plan or some other source. TRACKLINE creates a probability distribution about a base track. This track can be constructed from as many as 10 segments, each of which can be a portion of a rhumb line or of a great circle.

The motion of the target about each base track segment is specified by three circular normal probability distributions corresponding to target position at the initial, mid-point, and endpoint of each segment. Each simulated target track is obtained by drawing random numbers for target position from these distributions and then connecting the points with straight lines.

Figure 3 illustrates a typical situation. The target's point of departure and intended destination are assumed known, and a base track is constructed between these points. The base track might be taken from the target's float plan or hypothesized from the target's past habits. In the case illustrated by Figure 3, there are three track segments. The 50% circles of uncertainty are assumed to grow in size to about midway along the track and then diminish. Since the point of departure and intended destination are assumed to be known, the extreme endpoints of the entire track have zero uncertainty.

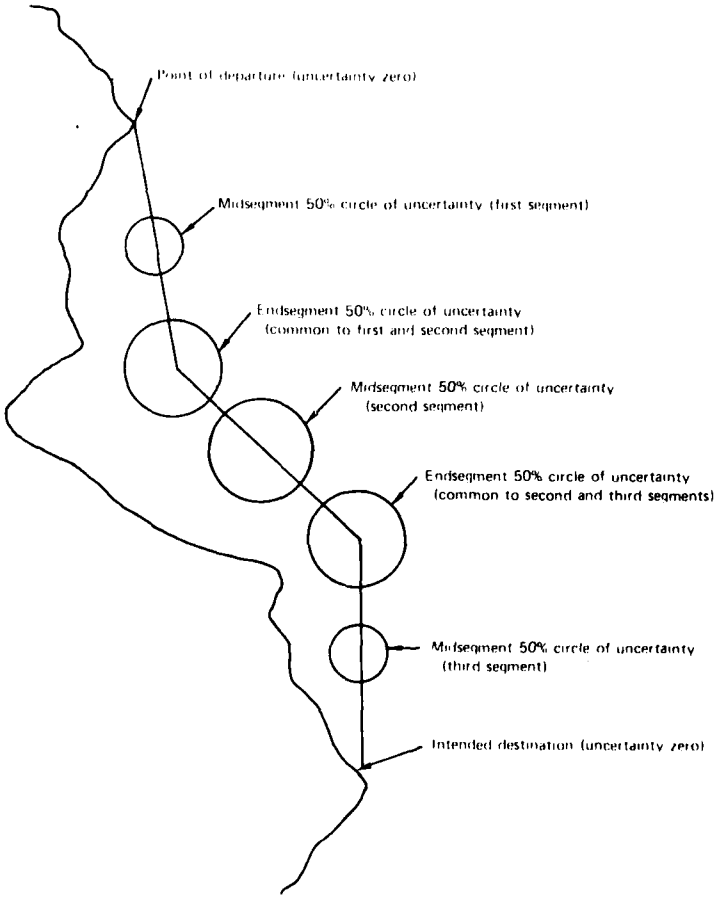


FIGURE 3 Description of trackline uncertainties

In some cases, there is information which leads one to suspect that the distress is more likely to have occurred on one part of the track than on another. For example, as mentioned above, the track may have passed through an area of storms and heavy seas. If desired, the target location probability distribution generated by TRACKLINE can be made to have a higher density in such an area. This is done by specifying the highest probability point along base track together with the odds that the distress occurred there rather than at the extreme endpoints of the track. These inputs determine a truncated triangular probability density for the fraction of track covered before the distress incident occurred.

Updating for Target Motion

The DRIFT program is used to alter a target location probability distribution to account for the effects of drift. Normally, the DRIFT program will cause the center of the distribution to move to a new location and the distribution to become more diffuse.

Target motion due to drift complicates the maritime search problem. The prediction of drift must account for the effects of both sea current due to prevailing circulation and predicted or observed surface wind. Any object floating free on the ocean surface is transported directly by surface current, and one component vector of drift is therefore equal to the predicted current vector. A statistical file collected from ship reports over many years has been assembled by the Coast Guard and arranged by geographical location and month of the year. The file in use in the CASP system covers most of the North Atlantic and North Pacific Oceans.

As mentioned above, wind is also important in predicting target motion. With regard to this factor, there are two major considerations. The first is the drift caused by the wind impinging on the drifting object's surface area above water; this is called "leeway." The speed and direction of leeway is different for different objects, and is usually difficult to predict.

The second wind consideration is the movement of the surface layer of the ocean itself; this is called "local wind current." It is one of the most complex and least understood phenomena in the entire drift process.

The primary data source for surface winds in the CASP system is the Navy's Fleet Numerical Weather Central in Monterey, California. Every twelve hours their computers generate a time series for hemispheric wind circulation; three of these time series are used to produce certain geographical blocks of wind data which are transmitted to the Coast Guard for use by CASP. All data are retained in the system for two to three months.

The process of applying the drift motion to update a CASP distribution is simple enough. First, a set of total drift vector probability distributions is computed for various geographical areas based upon estimates of sea current, leeway, and local wind current. Then for each target location replication, a random vector of net drift is drawn from the appropriate probability distribution and used to move the target forward a short time. The procedure is repeated until the entire update time is taken into account.

Updating for Negative Search Results

Once a search has actually been conducted, one of the two search update programs, RECTANGLE and PATH (depending upon the type of search), is run to revise the target location probabilities to account for unsuccessful search. The effect is to reduce the probabilities within the area searched, and to increase them outside.

Updating the target location probabilities for negative search results is carried out by an application of Bayes' theorem. Recall that the target triple file contains J records of the form (X_j, Y_j, Φ_j) for $1 \leq j \leq J$, where the pair (X_j, Y_j) represents target position, and Φ_j represents the probability that the target replication would not have been detected by the cumulative search effort under consideration. The overall cumulative probability of detection taking all simulated targets into account is called search effectiveness probability (SEP) and is computed by the formula

$$\text{SEP} = 1 - \sum_{1}^{J} \Phi_j / J.$$

Let C be a region in the search area, and let B_j denote the event, "target corresponds to the jth replication and is in region C." The posterior probability $\Lambda(C)$ that the target is located in C given search failure is computed using Bayes' theorem by

$$\begin{aligned}\Lambda(C) &= Pr\{\text{Target in } C \mid \text{Search failure}\} = \sum_{j=1}^{J} Pr\{B_j \mid \text{Search failure}\} \\ &= \sum_{1}^{J} Pr\{\text{Search failure} \mid B_j\}\, Pr\{B_j\} / Pr\{\text{Search failure}\} \\ &= \sum_{j \in I} \Phi_j / \sum_{j=1}^{J} \Phi_j,\end{aligned}$$

where $I = \{j : (X_j, Y_j) \in C\}$ denotes the set of indicies corresponding to target replications in C.

Now suppose that q_j denotes the probability of failing to detect the jth target replication during a particular update period. Using the independence assumption, the new individual cumulative failure probability Φ_j is computed by

$$\Phi_j = q_j \Phi_j',$$

where Φ' denotes the cumulative failure probability prior to the last increment of search.

The computation of the conditional failure probability q_j is carried out in CASP by use of a (M, β, σ)-detection model as described below. Recall (e.g., see Koopman [2]) that the "lateral range" between searcher and target (both with constant course and speed) is defined as the distance at closest point of approach. The "lateral range function" gives single sweep cumulative detection probability for a specified lateral range for a specified period of time. The integral of the lateral range function is called the "sweep width" of the sensor.

The CASP programs* are based upon the assumption that the lateral range function for the search unit is rectangular and is described by two parameters, M and β. Here M denotes the total width of the swept path, and β denotes the probability that the target would be detected for lateral ranges less than or equal to $M/2$. The sweep width W for the rectangular lateral range function described above is given by

$$W = \beta M.$$

Navigational uncertainties ("pattern error") are introduced into the detection model by assuming each sweep is a random parallel displacement from the intended sweep. The random

* A [illegible] provided [illegible] use an inverse cube lateral range function as defined in [2] together with search pattern [illegible].

displacements are assumed to be independent identically distributed normal random variables with zero mean and standard deviation. This model was introduced by R. K. Reber (e.g., see Reber [4]) and used extensively in certain Navy search analyses.

Rectangular lateral range functions are a useful way of approximating more complex lateral range functions. If the actual lateral range function has sweep width M and is *nonzero* over an interval of width M, then one may define β to be the average detection probability over the effective range of the sensor, i.e., $\beta = W/M$. Appendix A of [4] shows that replacement of the actual lateral range function by a rectangular lateral range function with average probability β usually does not lead to significant errors in the computed value of probability of detection for parallel path search. Cases where there is significant disagreement occur when the lateral range function is close to zero over a large part of its support.

Let G_σ denote the cumulative normal probability distribution function. Let (u, v) denote the target's position in a coordinate system where the origin is at the midpoint of a given sweep, and where the u-axis is parallel to the sweep and the v-axis is perpendicular to the sweep. Then for fixed M, β, and σ, the single sweep probability $p(u, v)$ of detecting the target is given by

$$(1) \qquad p(u,v) = \beta\left[G_\sigma\left(u + \frac{L}{2}\right) - G_\sigma\left(u - \frac{L}{2}\right)\right]\left[G_\sigma\left(v + \frac{M}{2}\right) - G_\sigma\left(v - \frac{M}{2}\right)\right],$$

where L denotes the length of the sweep.

If there are K search legs to be considered, and if (u_j^k, v_j^k) denotes the coordinates of the jth simulated target position relative to the kth search leg, then the failure probability q_j is given by

$$(2) \qquad q_j = \prod_{k=1}^{K} [1 - p(u_j^k, v_j^k)].$$

The application of these formulas in programs PATH and RECTANGLE can now be discussed.

Path. Program PATH is used to represent general search patterns constructed from straight track segments. For example, PATH can be used to compute detection probabilities for a circle diameter search where the search tracks are intended to cover a given circle by making repeated passes through its center. PATH makes direct use of (1) and (2).

Rectangle. Program RECTANGLE has been designed for the special case where a rectangle is searched using parallel sweeps. RECTANGLE reduces the computing time and amount of input that otherwise would be required using program PATH. For a point outside the designated rectangle, the probability of detection q_j is assumed to be 0. For a point inside the designated rectangle, "edge" effects are ignored and an average probability of detection is computed as if there were an infinite number of sweeps, each infinitely long.

The following line of reasoning originated with R. K. Reber. Reber [4] presents results in the form of curves and tables, and these have been adapted to program RECTANGLE by use of polynomial approximations. Let S denote the spacing between sweeps. Since the sweeps are assumed to be parallel and of infinite extent, the coordinate v_j^k expresses the lateral range for the kth sweep and the jth simulated target location and is given by

$$v_j^k = \mu_j + kS$$

for $-\infty < k < \infty$ and a number μ_j such that $|\mu_j| \leq S$.

Now for arbitrary μ, refer to (1) and (2) and define g by

$$(3)\quad g(\mu, S) = \prod_{k=-\infty}^{\infty} [1 - p(u, \mu + kS)] = \prod_{k=-\infty}^{\infty} \beta\left[G_{\sigma}\left(\mu + kS + \frac{M}{2}\right) - G_{\sigma}\left(\mu + kS - \frac{M}{2}\right)\right]$$

Note that since the sweeps are assumed to be of infinite length, one has $u = \infty$ and g defined by (3) does not depend upon u. The function g is periodic in its first argument with period μ. Let $\hat{g}(S)$ denote the average value of $g(\mu, s)$ with respect to the first argument. Then

$$\hat{g}(S) = \frac{1}{S}\int_0^S g(\mu, S)\, d\mu.$$

The function $\hat{g}$ has been tabulated in [4] and is used in program RECTANGLE to represent the failure probability $q_i = \hat{g}(S)$ for a point lying within the designated search rectangle. RECTANGLE and PATH agree (as they should) when PATH is used to represent a parallel path search.

Search Optimization

Two programs, MAP and MULTI, are used for optimizing the allocation of search effort. MAP provides a quick way of determining the search cells which should receive effort based upon a constraint on total track line miles available. MULTI determines search areas for multiple search units under the constraint that each unit must be assigned a uniform coverage of a rectangle and that the rectangles for the various search units do not overlap.

The method used in program MAP is based upon use of an exponential detection function (see Stone [8]) introduced by Koopman [3] and does not impose constraints on the type of search pattern employed. The primary usefulness of this program is to provide the search planner with a quick method for defining the area of search concentration. The following paragraphs give a brief sketch of the methods used in these optimization programs.

Map Let there be N search cells, and for $1 \leqslant n \leqslant N$ let p_n and α_n denote, respectively, the target location probability and the area associated with the nth cell. The probability *density* for target location in the nth cell is given by $d_n = p_n/\alpha_n$. Suppose that total search effort is measured by the product of track line miles and sweep width.

Let γ denote an allocation of search effort where $\gamma(n)$ denotes the amount of search effort (measured in area swept) allocated to the nth cell. Probability of detection $P_D[\gamma]$ is computed using an exponential effectiveness function, i.e.,

$$P_D[\gamma] = \sum_{n=1}^{N} p_n[1 - \exp(-\gamma(n)/\alpha_n)].$$

The objective is to maximize P_D subject to a constraint on total effort available. This is easily done using the techniques introduced by Koopman [3]; easier proofs are provided in Stone [8] and Wagner [12].

It can be shown that under the above assumptions, the initial increments of effort should be concentrated in the highest probability density cells, and that there should be a succession of expansions to cells having lower target location probability density.

In order to derive the formulas used in program MAP, a new collection of equi-density search regions is formed made up of the unions of all cells having equal probability density. Let

K = the number of equi-density regions
d_k = the probability density for region k
I_k = the set of indices corresponding to the cells comprising region k
A_k = the area of region k.

Using the above notation

$$A_k = \sum_{i \in I_k} \alpha_i.$$

Let E_k denote the total effort which must be expended *before* the optimal search expands into the kth region. Assume that the equi-density regions have been ordered beginning with the region having the highest density. Since search begins in the highest density region, we have $E_1 = 0$. It can be shown that in general for $k \geq 2$

$$E_k = E_{k-1} + (\ln d_{k-1} - \ln d_k) \sum_{m=1}^{k-1} A_m. \tag{4}$$

Figure 4 shows output from program MAP illustrating the use of (4). The list shows the 25 highest probability cells specified by the latitude and longitude of the southeast corner. Each cell is 15 minutes wide, and the numbers in the last column correspond to the values E_k given by (4). The planning advice given in [10] is to apply search effort to any cell for which the value in the effort column is less than the total effort available.

TOP 25 PROBABILITY		LOCATION	(S.E. CORNER)	EFFORT
1	0.05133	43-0N	69-45W	
2	0.04167	42-45N	69-30W	35.0
3	0.04133	43-0N	70-0W	36.3
4	0.04100	43-0N	69-30W	40.3
5	0.03567	43-15N	69-30W	129.4
6	0.03467	43-15N	69-45W	152.8
7	0.03333	42-45N	69-15W	199.6
8	0.03267	42-30N	69-15W	227.5
9	0.03267	42-45N	69-45W	222.2
10	0.03267	43-15N	70-0W	210.1
11	0.03200	42-30N	69-30W	264.1
12	0.02800	43-0N	69-15W	491.5
13	0.02733	42-45N	70-0W	547.2
14	0.02533	43-0N	70-15W	701.3
15	0.02267	42-30N	69-0W	976.5
16	0.02233	43-15N	69-15W	983.1
17	0.02167	42-30N	69-45W	1095.1
18	0.02133	43-15N	70-15W	1104.4
19	0.02100	42-45N	69-0W	1175.3
20	0.01867	43-30N	69-30W	1505.8
21	0.01867	43-30N	69-45W	1505.8
22	0.01800	43-0N	69-0W	1659.9
23	0.01667	42-30N	68-45W	1968.0
24	0.01600	42-15N	69-0W	2137.7
25	0.01600	42-15N	69-15W	2137.7

FIGURE 4. Optimal allocation of effort produced by MAP

Notice that the numbers in the effort column are not necessarily increasing. This is because the list is ordered according to containment probability rather than probability density.

Multi. As mentioned above, program MAP does not take into account "simplicity" constraints which are considered important in operational planning. Program MULTI was designed to overcome this drawback in cases where multiple search units are deployed in the same search area.

The first simplicity constraint introduced is that each unit will be assigned to uniformly search a rectangle. Figure 5 shows the dimensions of the optimal rectangle and the resulting probability of detection under the assumption that the target location probability distribution is normal. In order to use this figure, one first computes the normalized effort E^* by the formula

$$E^* = \frac{RT}{\sigma_{\max}\sigma_{\min}},$$

where R is the sweep rate of the unit, T is the total search time, and $\sigma_{\max}$ and $\sigma_{\min}$ are the standard deviations of the normal distribution when referred to principal axes. The optimal search rectangle will have half side given by $U^*\sigma_{\max}$ and $U^*\sigma_{\min}$ where the size factor U^* is given by the designated curve with values read along the outer vertical scale.

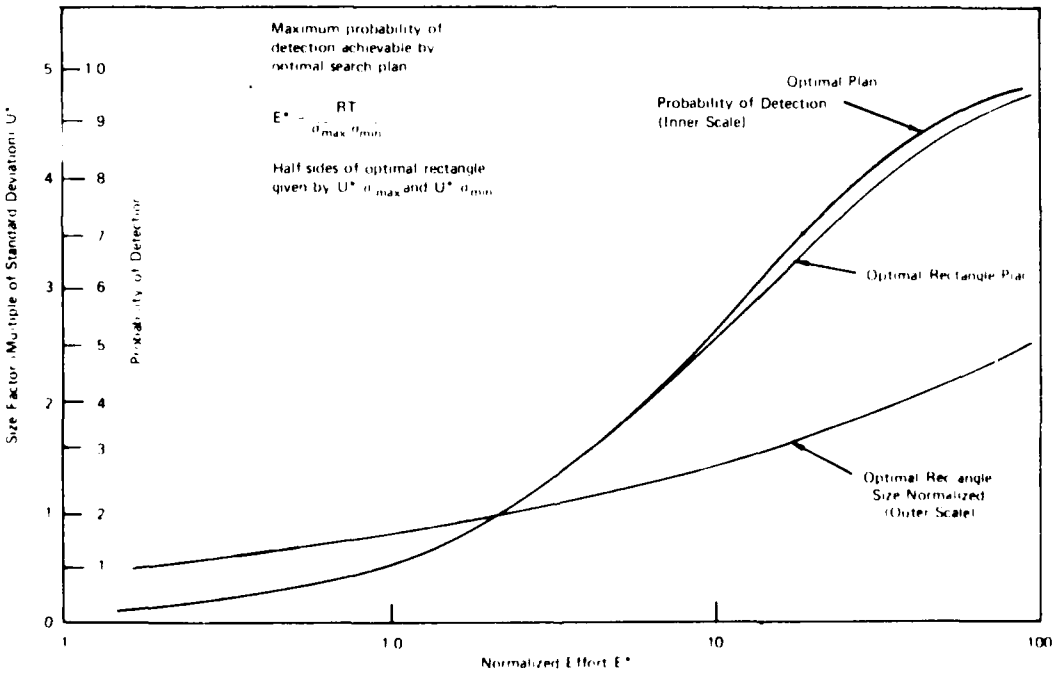


FIGURE 5. Optimal search rectangle

Figure 5 provides curves to determine the probability of detection for the optimal rectangle plan and for the unconstrained optimal plan. It is interesting to note that in all cases the probability of detection provided by the optimal rectangle plan is at least 95% of that provided

by the unconstrained optimal plan. Thus, under the assumption stated, uniform search of the optimal rectangle can be recommended without hesitation since, in most cases, the simplicity of the rectangle plan is more important than the small improvement in effectiveness obtained by the more complicated optimal plan.

MULTI is capable of allocating the effort of up to 5 search units to nonoverlapping rectangles in a way which is intended to maximize overall probability of detection. The first step in this procedure is to approximate the target location probability distribution by the weighted average of k bivariate normal distribution where $i \leqslant k \leqslant 3$. This is done by locating the three highest local maxima in the smoothed cell distribution and then associating each simulated target position with the nearest cluster point. If three local maxima cannot be found, then the procedure is carried out with one or two local maxima. The mean and covariance matrix of each cluster are calculated to determine the parameters of the approximating normal distribution.

The program next considers all possible assignments of search units to one of the three approximating probability distributions. Since there are a maximum of five units and three distributions, there are at most $3^5 = 243$ different ways of assigning units to distributions. For each assignment, the program sums up the total effort available to search each distribution and then computes the resulting optimal rectangle and associated probability of detection. If P_k denotes the conditional probability of detecting the target with optimal rectangle search given that the target has the kth distribution $(1 \leqslant k \leqslant k)$, then probability of detection Δ for the allocation is given by

$$\Delta = \sum_{k=1}^{K} P_k D_k.$$

The program prints the allocation which gives the maximum probability of detection and notes whether any of the rectangles overlap. If overlap occurs, then the next ranking allocation is printed, and so on. This continues until an allocation without overlap is found or until the top five allocations have been listed together with their associated detection probabilities. Finally, when several units are assigned to the same rectangle, it is subdivided in a way which preserves the uniform coverage.

Recently an alternative method for multiple unit allocation has been developed (see Discenza [1]) based upon integer programming considerations.

3. CASP CASE EXAMPLE

On 12 September 1976 the sailing vessel S/V Spirit departed Honolulu enroute San Francisco Bay. The owner, who was awaiting its arrival in San Francisco, reported concern for the vessel to the Coast Guard on 14 October 1976 after it had failed to arrive. An Urgent Marine Information Broadcast (UMIB) was initiated on 17 October. The following day, a merchant vessel the M/V Oriental Financier reported recovering a life raft with two survivors from the S/V Spirit which had sunk in heavy seas in mid-Pacific on the morning of 27 September. Survivors indicated three more crewmembers in a separate raft were still adrift. This information opened an extensive six day air and surface search for the missing raft that eventually located the raft with one of the missing persons on board.

Each day's search was planned utilizing computer SAR programs. Initial distress position information was gained by radio-telephone debriefing of the survivors aboard the M/B Oriental Financier on several occasions. The search began 19 October based on a SARP* datum for a raft without a drogue from an initial reported position of 36N 136W. The second day's search was based on a SARP datum for a position 160 nautical miles to the northeast from the previous position (this position being determined from further debriefing of the survivors over radio-telephone). The third through the six days' searches were planned utilizing CASP output from a POSITION scenario consisting of an ellipse with a 160 mile major axis and a 60 mile minor axis. The CASP program was updated by RECTANGLE and DRIFT daily, and search areas assigned to cover the highest cells which could be reached taking into account search unit speed and endurance.

The following chronology is based upon the official USCG report and describes the utilization of CASP in the search planning. This case is a good illustration of the many uncertainties which must be analyzed during a search and the way both negative and positive information contribute to eventual success.

21 October 1976

Search planning for the day's operations utilized the CASP program for the first time. New probable distress position information given by the survivors was evaluated and the CASP program was initiated using a POSITION scenario with center length 160 miles and width 60 miles oriented on 046°T, with the southwest end at position 36N 136W. This scenario was to be used for the rest of the search. A search plan was generated for the 21 October search covering approximately 8 of the 10 highest CASP cells as given in MAP. Ten units were designated for the day's efforts and consisted of 3 Coast Guard, 2 Navy, and 4 Air Force aircraft and the USS Cook.

The first aircraft which arrived on scene for the day's search reported the weather in the search area as ceiling varying 200-1500 feet (scattered), wind from 330° at 8 knots, seas 4 feet, and visibility unlimited except in occasional rain showers.

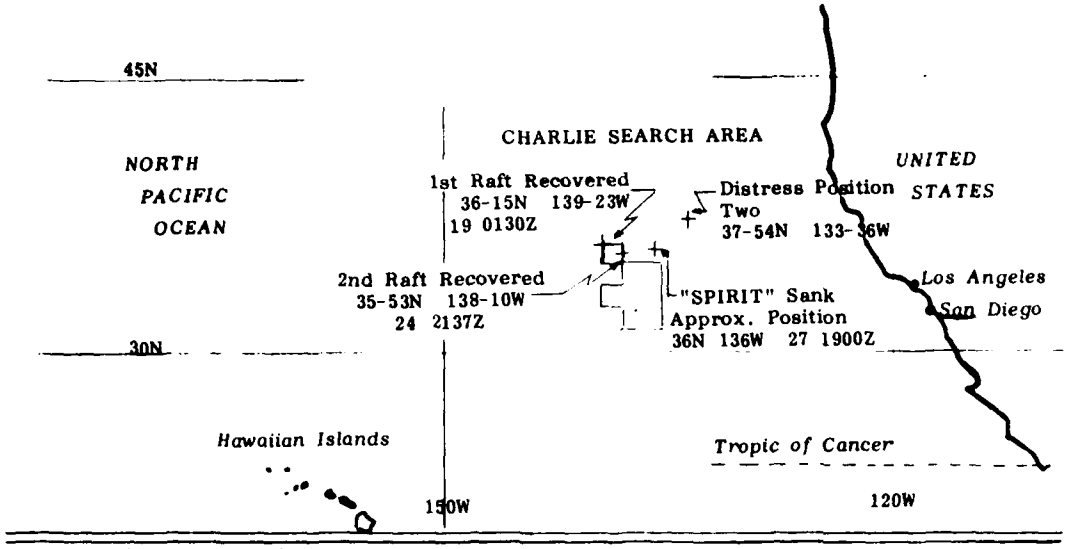
At 3:06 PM an aircraft located what appeared to be the life raft of recovered survivors in position 35-38N 138-12W. M/V Oriental Financier had been unable to recover this raft when the survivors were rescued. The USS Cook investigated and reported negative results.

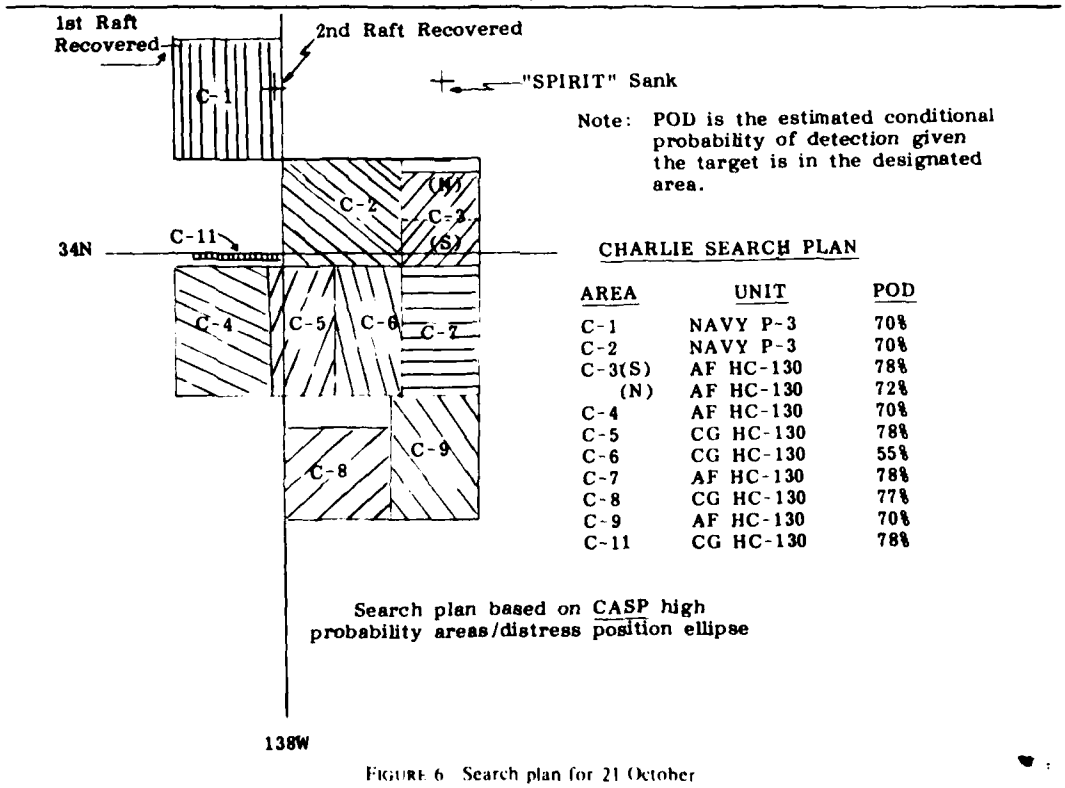
Figure 6 shows the search plan for 21 October. Note that the target was eventually found on 24 October in the first designated area C-1. There is, of course, no way of knowing where the target was on the 21st.

22 October 1976

Planning for day's search was done using updates from the CASP program. Search units, consisting of 17 aircraft (3 Coast Guard, 6 Navy, and 8 Air Force) and the USS Cook, were designated areas totaling 67,920 square miles for the day's effort. Areas assigned were determined from the MAP's twelve highest cells. High altitude photographic reconnaissance flight utilizing U-2 aircraft was also scheduled, cloud coverage permitting, to cover an area of 57,600 square miles.

*A computer program implementing methods described in the National SAR Manual and a precursor to CASP





Note: POD is the estimated conditional probability of detection given the target is in the designated area.

CHARLIE SEARCH PLAN

AREA	UNIT	POD
C-1	NAVY P-3	70%
C-2	NAVY P-3	70%
C-3(S)	AF HC-130	78%
(N)	AF HC-130	72%
C-4	AF HC-130	70%
C-5	CG HC-130	78%
C-6	CG HC-130	55%
C-7	AF HC-130	78%
C-8	CG HC-130	77%
C-9	AF HC-130	70%
C-11	CG HC-130	78%

Search plan based on CASP high probability areas/distress position ellipse

FIGURE 6 Search plan for 21 October

The first aircraft on scene for the day's search reported the weather in the general area as ceiling 1800 feet (broken), winds from 150° at 6 knots, seas 2 feet, and visibility 15 miles.

Search conducted during daylight hours utilized 15 aircraft, the USS Cook, and a U2 high altitude reconnaissance flight. The USS Cook was unable to relocate debris sighted during previous day's search. Two Air Force aircraft failed to arrive on scene prior to darkness and were released. Aircraft on scene searched 88 percent of 67,920 square miles assigned and obtained POD's ranging from 50 to 82 percent. The high altitude photographic reconnaissance flight was conducted from an altitude of approximately 50,000 feet.

The CGC Campbell arrived on scene and relieved the USS Cook.

23 October 1976

The Rescue Coordination Center (RCC) was advised by the Air Force that development of high altitude film had shown an "orange dot" in position 35-16N 139-05W. The photographed object was described as a round orange object, approximately 7 feet in diameter, floating on the surface of the water.

Search planning was done using updates from the CASP program. Search units, consisting of the CGC Campbell and 8 aircraft (2 Coast Guard, 3 Navy, and 3 Air Force), were assigned areas of highest CASP cells. The object photographed by reconnaissance aircraft was drifted by SARP and the CGC Campbell and 1 aircraft dedicated to locate it.

The first aircraft on scene for the day's search reported weather in the search area as ceiling 2000 feet, wind from 200° at 12 knots, seas 2 feet, and visibility 15 miles.

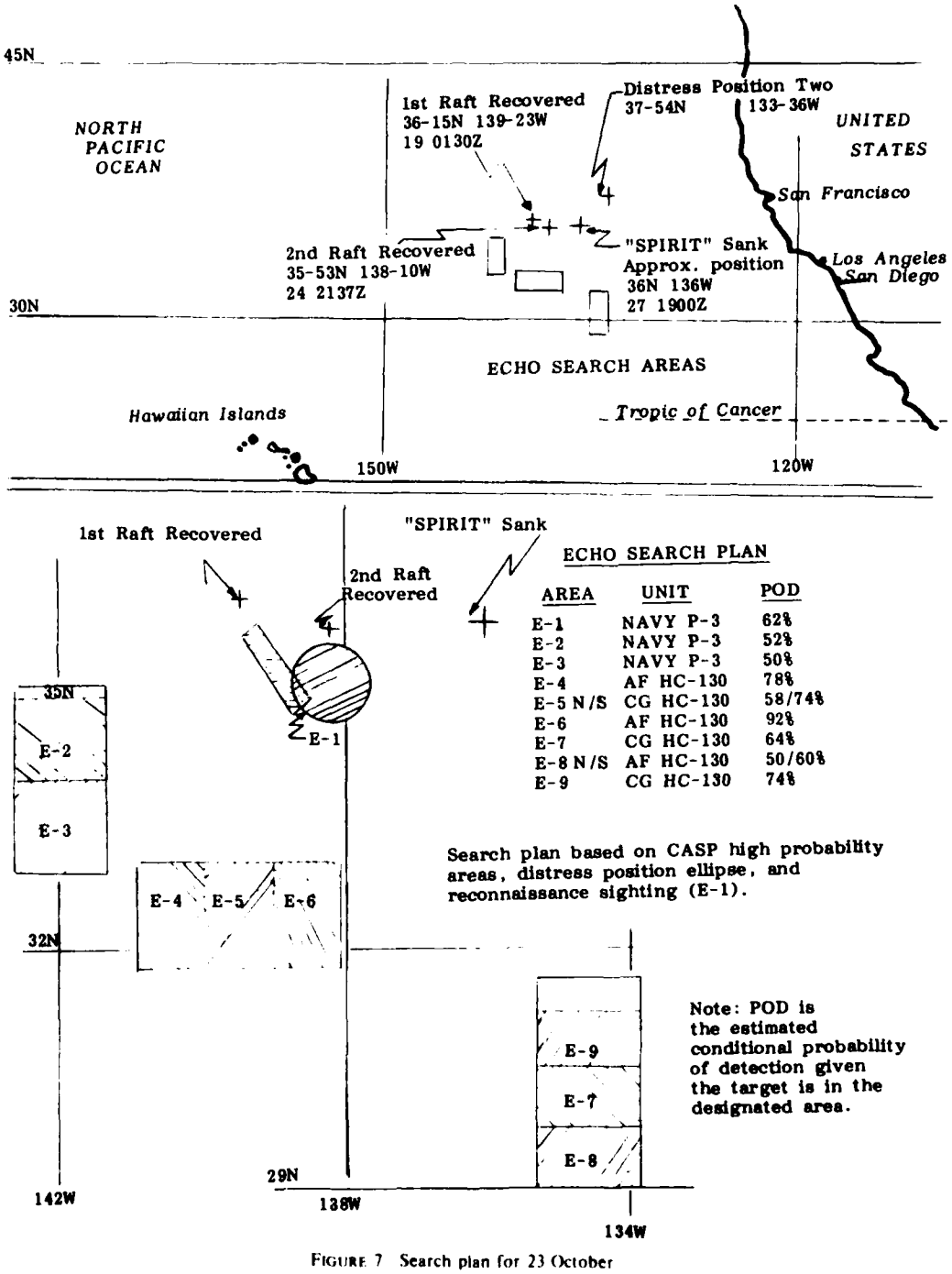
Search conducted during daylight hours utilized 8 aircraft and CGC Campbell. Search units covered 97 percent of the assigned 34,300 square miles with POD's ranging from 50 to 92 percent. Several sightings of assorted flotsam were reported but none linked to Spirit or rafts. The object photographed by the high altitude reconnaissance flight on 22 October was not relocated by search units.

Figure 7 shows the search plan for 23 October. Although not indicated in the chart, the position where the target was found on the 24th is in the second highest probability density cell from the CASP map.

24 October 1976

Search planning for the day's operations was done using updates from the CASP program. Search units consisting of the CGC Campbell and 5 aircraft (2 Coast Guard and 3 Navy) were assigned areas of highest CASP probability totaling 18,082 square miles, with CGC Campbell and one Coast Guard aircraft designated for location of the object reported by the reconnaissance flight.

The position of the reconnaissance flight sighting of 22 October was drifted utilizing SARP and the new position passed to CGC Campbell for search purposes. The 11:00 AM SARP datum was computed to be 35-29.4N 138-39.2W with standard first search radius of 16.9 miles. The search plan is shown in Figure 8.



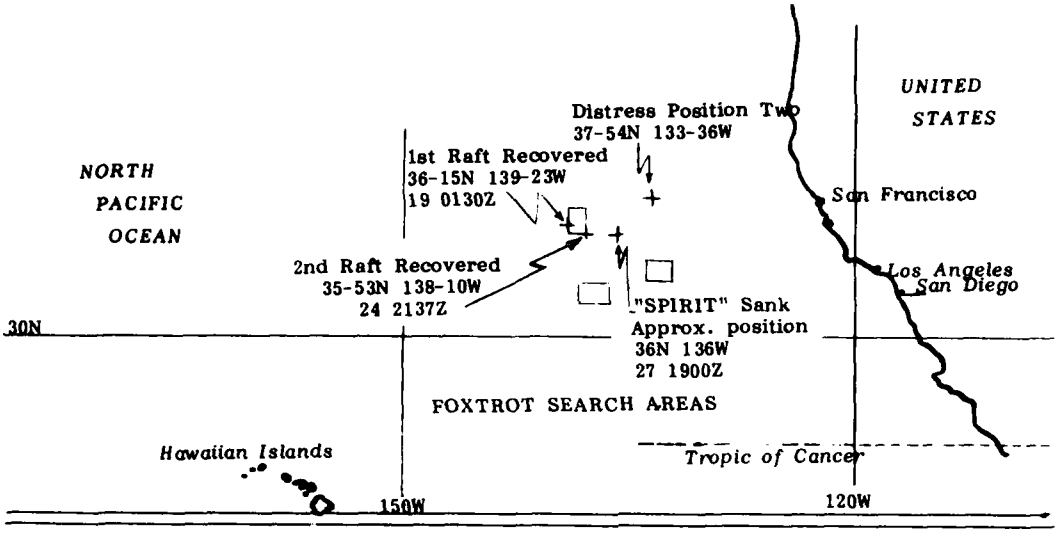
ECHO SEARCH PLAN

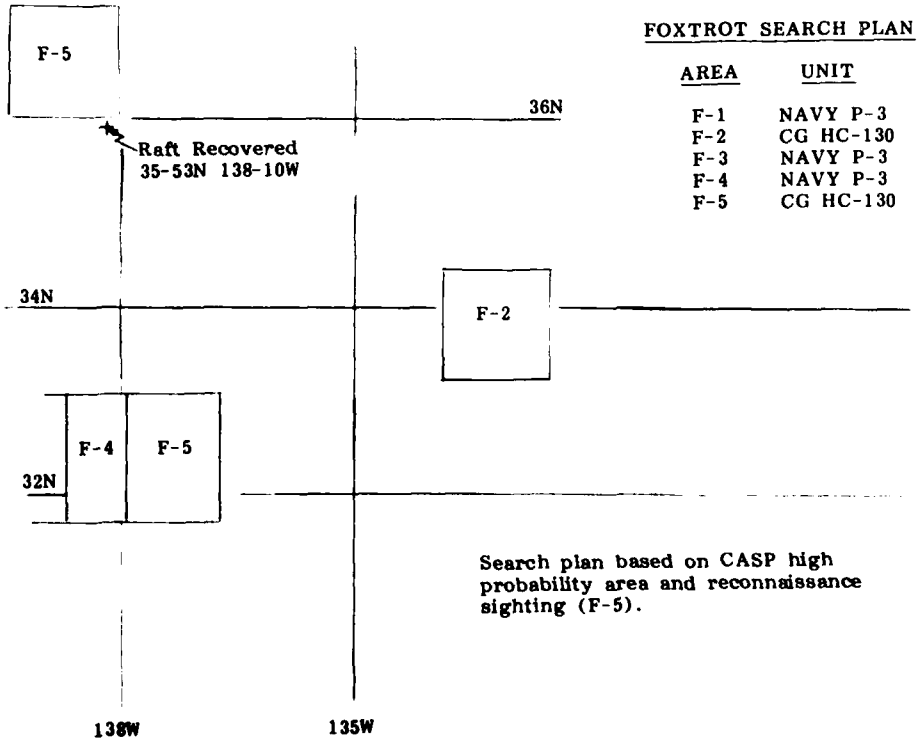
AREA	UNIT	POD
E-1	NAVY P-3	62%
E-2	NAVY P-3	52%
E-3	NAVY P-3	50%
E-4	AF HC-130	78%
E-5 N/S	CG HC-130	58/74%
E-6	AF HC-130	92%
E-7	CG HC-130	64%
E-8 N/S	AF HC-130	50/60%
E-9	CG HC-130	74%

Search plan based on CASP high probability areas, distress position ellipse, and reconnaissance sighting (E-1).

Note: POD is the estimated conditional probability of detection given the target is in the designated area.

FIGURE 7 Search plan for 23 October





FOXTROT SEARCH PLAN

AREA	UNIT
F-1	NAVY P-3
F-2	CG HC-130
F-3	NAVY P-3
F-4	NAVY P-3
F-5	CG HC-130

Search plan based on CASP high probability area and reconnaissance sighting (F-5).

FIGURE 8. Search plan for 24 October

The first aircraft on scene for the day's search reported weather in the search area as ceiling 1500 feet, wind from 000° at 7 knots, seas 3 feet, and visibility 10 miles.

The CGC Campbell reported locating a rusty, barnacle encrusted 55 gallon drum in position 35-27.2N 138-39.0W.

At 12:05 PM the search met with success! A Coast Guard HC-130H reported sighting a raft in position 36-03N 138-00W with at least one person on board. The CGC Campbell proceded enroute to investigate, and at 2:37 PM CGC Campbell reported on scene with the raft in position 35-53N 138-10W. A small boat was lowered to recover the survivor, and at 3:01 PM all search units were released from the scene.

4. TRAINING

CASP training began with an operational testing phase in cooperation with the New York RCC. This operational testing was useful in orienting the personnel to the benefits derived from more detailed search planning, and provided an idea of what the full training problem was going to be like.

Coincident with this, a training manual [9] and a completely new combined operating handbook [10] were developed encompassing all of the operational computer services available.

At the time of official implementation in February 1974, a special four-day class was conducted in the operation of the CASP system; this class was attended by one representative from each Rescue Coordination Center. It was intended that these persons would learn the system thoroughly and return to their respective commands and teach others. This plan was marginally successful, and worked only in those cases where an extremely capable individual was selected for attendance.

During the next six months, personnel from the Operations Analysis Branch visited each East Coast RCC for one week apiece in order to provide additional training. Subsequently, the same visit schedule was repeated on the West Coast.

Another valuable tool for training has been telephone consultation. Fortunately, all messages into and out of the computer are monitored at New York, and personnel can be helped with the details of input and output with a quick telephone call on the spot.

Finally, the National Search and Rescue School has made CASP training a regular part of its curriculum. The school, located on Governors Island, is responsible for initial training of all RCC personnel (among many others) in the techniques of search and rescue. The present SAR school training session is four weeks in duration with the fourth week devoted to computer search planning systems training. Over half of this time is devoted directly to CASP.

The Coast Guard is currently in the process of separating its administrative and operational systems by establishing an Operational Computer Center. This new Center will give rescue coordinators direct access to CASP through on-line terminals and will improve CASP's availability and reliability. Interactive program control will make the modules easier to use.

The application of CASP in operational situations has been quite successful, in spite of significant encumberances associated with computer and communications services.

Continued oceanographic research programs will expand CASP's applicability to important in-shore regions. Implementation of the new multi-unit allocation algorithm [1] is expected to simplify the search area assignment problem. These additional capabilities coupled with improved computer access and reliability should make CASP an even more valuable planning tool in the future.

ACKNOWLEDGMENTS

The development, implementation, training, and utilization of CASP represents the contributions of individuals far too numerous to mention by name in this paper. Foremost among these are the officers and men who use CASP in the RCCs and without whom the system would be useless. The contributions of the following individuals to the support and development of CASP are specifically acknowledged: C. J. Glass, R. C. Powell, G. Seaman, V. Banowitz, F. Mittricker, R. M. Larrabee, J. White, J. H. Hanna, L. D. Stone, D. C. Bossard, B. D. Wenocur, E. P. Loane, and C. A. Persinger.

REFERENCES

[1] Discenza, J.H., "Optimal Search with Multiple Rectangular Search Areas," Doctoral Thesis, Graduate School of Business Administration, New York University (1979).

[2] Koopman, B.O., "The Theory of Search, Part II, Target Detection," Operations Research, *4*, 503-531 (1956).

[3] Koopman, B.O., "The Theory of Search, Part III, The Optimum Distribution of Searching Effort," Operations Research, *5*, 613-626 (1957).

[4] Reber, R.K., "A Theoretical Evaluation of Various Search/Salvage Procedures for Use with Narrow-Path Locators, Part I, Area and Channel Searching," Bureau of Ships, Minesweeping Branch Technical Report, No. 117 (AD 881408) (1956).

[5] Richardson, H.R., Operations Analysis, February (1967). Chapter V, Part 2 of *Aircraft Salvage Operation, Mediterranean*, Report to the Chief of Naval Operations prepared by Ocean Systems, Inc. for the Supervisor of Salvage and the Deep Submergence Systems Project.

[6] Richardson, H.R. and L.D. Stone, "Operations Analysis During the Underwater Search for Scorpion," Naval Research Logistics Quarterly, *18*, 141-157 (1971).

[7] Shreider, Yu. A., *The Monte Carlo Method* (Pergamon Press, 1966).

[8] Stone, L.D., *Theory of Optimal Search* (Academic Press, 1975).

[9] U. S. Coast Guard, Commander, Atlantic Area, CASP Training Course, 19-22 February (1974).

[10] U. S. Coast Guard, Computerized Search and Rescue Systems Handbook (1974).

[11] U. S. Coast Guard, National Search and Rescue Manual (1970).

[12] Wagner, D.H. "Nonlinear Functional Versions of the Neyman-Pearson Lemma," SIAM Review, *11*, 52-65 (1969).

CONCENTRATED FIRING IN MANY-VERSUS-MANY DUELS

A. Zinger

University of Quebec at Montreal
Montreal, Canada

ABSTRACT

A simple stochastic-duel model, based on alternate firing, is proposed. This model is shown to be asymptotically equivalent, for small hit probabilities, to other known models, such as simple and square duels. Alternate firing introduces an interaction between opponents and allows one to consider multiple duels. Conditions under which concentrated firing is better or worse than parallel firing are found by calculation and sometimes by simulation. The only parameters considered are the combat group sizes (all units within a group are assumed identical), the hit probabilities and the number of hits necessary to destroy an opposing unit.

1. INTRODUCTION

Two extremes for the modeling combat attrition are given by the so-called Lanchester theory of combat, which treats combat attrition at a *macroscopic* level, and by the theory of stochastic duels, which treats combat attrition at a *microscopic* level and considers individual firers, target acquisition, the firing of each and every round, etc. (see Ancker [1, pp. 388-389] for further details). Actual combat operations are, of course, much more complex than their representation by such relatively simple attrition models and may also be investigated by means of much more detailed Monte Carlo combat simulations. Unfortunately, such detailed Monte Carlo simulations usually fail to provide any direct insights into the dynamics of combat without a prohibitive amount of computational effort. In the paper at hand, we will consider a relatively simple stochastic-duel model to develop some important insights into a persisting issue of military tactics (namely, what are the conditions under which concentration of fire is "beneficial").

In his now classic 1914 paper, F.W. Lanchester [10] (see also [11]) used a simple deterministic differential-equation model to quantitatively justify the principle of concentration, i.e., a commander should always concentrate as many men and means of battle at the decisive point. From his simple macroscopic model, Lanchester concluded that the "advantage shown to accrue from fire concentration as exemplified by the n square law is overwhelming." However, this conclusion depends in an essential way on the macroscopic differential-equation attrition model used by Lanchester [10], [11] (see Taylor [14] for further discussion) and need not hold for microscopic stochastic-duel models of combat attrition. In fact, this paper shows that for such microscopic duel models it is not always "best" to concentrate fire.

Subsequently, many investigators have commented on the benefits to be gained from concentrating fire. For example, in his determination of the probability of winning for a stochastic analogue of Lanchester's original model, Brown [6] stressed the fact that the model applied to

cases of concentrated firing by both sides. Other investigators of deterministic Lanchester-type models from the macroscopic combat-analysis point of view have also stressed this point (e.g. see Dolansky [7], Taylor [13], and Taylor and Parry [15]). Recently, Taylor [14] has examined the decision to initially commit forces in combat between two homogeneous forces modeled by very general deterministic Lanchester-type equations. He showed that it is not always "best" to commit as much as possible to battle initially but that the optimal decision for the initial commitment of forces depends on a number of factors, the key of which is how the trading of casualties depends on the victor's force level and time.

The first reference to problems of strategy in multiple duels is found in Ancker and Williams [2], who study the case of a square duel (2 vs 2) and arrive at the right conclusion that parallel firing is better than concentrated firing. This is a natural conclusion since only one hit is necessary to achieve destruction, and in concentrated firing there is a certain amount of over-killing. In 1967, Ancker [1] makes suggestions for future research concerning mutliple duels and states explicitly that the difficulties lie in the strong interaction between the contestants. The possibility of needing more than one hit to achieve destruction in the simple duel situation was introduced by Bhashyam [4] in 1970.

The purpose of this paper is to combine some of the above mentioned concepts, in order to gain insight concerning a problem of strategy in multiple duels—should one concentrate one's fire or not?

2. ASSUMPTIONS AND NOTATION

Let us consider two forces A and B that meet each other in combat. A consists of M units and B of N units.

The following assumptions are made:

1. Firing is alternating, volley after volley, i.e., A fires all weapons simultaneously, then B and so on until all units of a force are destroyed. This is contrary to the usual assumption of either simultaneous firing or random firing within some time intervals as found in Robertson [12], Williams [17], Helmbold [8], [9], Thompson [16], Ancker [3]. It is felt, and will be shown in a few cases, that for relatively small probabilities of hitting, this approach gives results comparable to Ancker and Williams [2]. We will denote by V_{ij} the probability of i winning if j shoots first $i, j = A,B$. The unconditional probability of winning will be denoted by V_A or V_B.

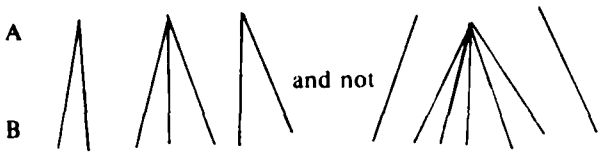
2. Hit probabilities are constant and are respectively p_A and p_B, with $q_i = 1 - p_i$, $i = A,B$.

3. Each unit of force A requires K_A hits to be destroyed. Same for B and K_B.

4. The supply of ammunition is unlimited.

5. There is no time limit to score a hit.

6. In a multiple duel (more than 1 vs 1) the units of A concentrate their fire on a single unit of B while the units of B each fire at a different unit of A, or spread their fire over all available units of A, this last case occurs when $M < N$. B has to allow an amount of concentration in order not to lose some shots. Concentration will be kept at a minimum to preserve as much parallelism as possible. For example if $M = 3$ and $N = 7$ the pattern of fire for B has to be



7. The most general notation, for example, $V_{A|B}(M,N,K_A,K_B,p_A,p_B)$ will be avoided if possible and replaced by an appropriate simpler form.

Before proceeding, a general remark ought to be made: most of the difficulties come from the asymmetry in the situation and from the interaction between the opponents. The same model has to express concentration, dispersion and partial concentration of fire. Moreover, the probability of winning depends upon the whole past history of the duel.

3. MULTIPLE DUEL. ONE HIT SUFFICIENT TO DESTROY

Let $K_A = K_B = 1$ and let $E(i,j)$ be the state of group A with i units, and of group B with j units.

If A fires first, the next state is

$E(i,j)$ with probability q_A^j and

$E(i,j-1)$ with probability $1 - q_A^j$.

When B fires, let us first consider the case when $j < i$. Then,

$E(i,j)$ becomes $E(i-k,j)$, $k = 0, \ldots, j$ with probability $\binom{j}{k} p_B^k q_B^{j-k}$

and

$E(i,j-1)$ becomes $E(i-k, j-1)$, $k = 0, \ldots, j-1$ with probability $\binom{j-1}{k} p_B^k q_B^{j-1-k}$.

If on the other hand $j \geqslant i$ some regrouping has to be done.

Let $j = ai + b$ with $b < i$, $a, b \in I^+$. The regrouping which spreads the fire the most is given by

a shots are fired with a probability of success

$1 - (1 - p_B)^a = 1 - A_0$ at each of $i - b$ targets

$a + 1$ shots are fired with a probability of success

$1 - (1 - p_B)^{a+1} = 1 - A_1$ at each of b targets.

Define $r = \min(i,j)$. Then both cases $j < i$ and $j \geqslant i$ are identical if one defines the probability of transition from state $E(i,j)$ to state $E(i-k, j)$ when B fires as

$$(3.1)\quad \theta(i,j,k,p_B) = \sum_{\substack{k_0+k_1=k \\ k_0=0,1,\ldots,r-b \\ k_1=0,1,\ldots,b}} \binom{b}{k_1}\binom{r-b}{k_0} (1-A_1)^{k_1}\, A_1^{\,b-k_1} (1-A_0)^{k_0}\, A_0^{\,r-b-k_0}$$

In the case $j < i$, $a = 0$, $k_0 = 0$ and $k_1 = k$.

It follows that if A starts and B returns fire once, the intial state $E(i, j)$ can become

$$E(i, j) \text{ with probability } q_A^i\, \emptyset\,(i, j, 0, p_B) = q_A^i\, q_B^j$$

$$E(i-k, j) \text{ with probability } q_A^i \emptyset\,(i, j, k, p_B),\ k = 1, \ldots, r$$

$$E(i-k, j-1) \text{ with probability } (1 - q_A^i)\ \emptyset\,(i, j-1, k, p_B),\ k = 0, 1, \ldots, r'$$

$$\text{where } r' = \min(i, j-1).$$

If B starts, the initial state $E(i, j)$ can become

$$E(i, j) \text{ with probability } q_A^i\ \emptyset\,(i, j, 0, p_B) = q_A^i\, q_B^j$$

$$E(i-k, j) \text{ with probability } q_A^{i-k}\ \emptyset(i,j,k,p_B),\ k = 1, \ldots, r$$

$$E(i-k, j-1) \text{ with probability } (1 - q_A^{i-k})\ \emptyset\,(i, j, k, p_B),\ k = 0, 1, \ldots, r''$$

$$\text{where } r'' = \min(i-1, j).$$

Let $V_{B|A}(M, N)$ denote the probability that group B wins with initial state $E(M, N)$ and A starts firing. Then

$$\begin{aligned} V_{B|A}(M, N) &= q_A^M\, q_B^N\, V_{B|A}(M, N) \\ &\quad + q_A^M \sum_{k=1}^{r} \emptyset\,(M, N, k, p_B)\ V_{B|A}(M-k, N) \\ &\quad + (1 - q_A^M) \sum_{k=0}^{r} \emptyset\,(M, N-1, k, p_B)\ V_{B|A}(M-k, N-1). \end{aligned} \tag{3.2}$$

This corresponds to a decomposition into all the mutually exclusive and exhaustive ways for B to win if A fires once and then B returns fire.

In a similar way

$$\begin{aligned} V_{B|B}(M, N) &= q_A^M\, q_B^N\, V_{B|B}(M, N) \\ &\quad + \sum_{k=1}^{r} q_A^{M-k}\ \emptyset\,(M, N, k, p_B)\ V_{B|B}(M-k, N) \\ &\quad + \sum_{k=0}^{r} (1 - q_A^{M-k})\ \emptyset(M, N, k, p_B)\ V_{B|B}(M-k, N-1). \end{aligned} \tag{3.3}$$

Since we have

$$V_{B|A}(M, 0) = V_{B|B}(M, 0) = 0 \qquad \text{all } M$$

and

$$V_{B|A}(0, N) = V_{B|B}(0, N) = 1 \qquad \text{all } N$$

we can calculate in succession all required probabilities. For example, since $\emptyset\,(1, 1, 1, p_B) = p_B$, one finds $V_{B|A}(1, 1) = q_A p_B / (1 - q_A q_B)$. Using $V_{B|A}(1, 1)$ and $\emptyset(1, 1, 0, p_B) = q_B$, $\emptyset(1, 2, 1, p_B) = 1 - q_B^2$, one finds $V_{B|A}(1, 2)$.

Explicitly, one gets, by assuming that A starts half the time,

$$\begin{aligned} V_B(M, 1) &= \frac{1}{2}\,((V_{B|A}(M, 1) + V_{B|B}(M, 1)) \\ &= \frac{1}{2}\, p_B^M\, q_A^{M(M-1)/2}\,(1 + q_A^M) \Big/ \prod_{i=1}^{M} (1 - q_B\, q_A^i). \end{aligned}$$

One can also obtain for $q_A = q_B = q$

$$V_B(2, 2) = \frac{1 + 4q + 4q^2 + 7q^3 + 4q^4 + 3q^5 + q^6}{2(1 + q)^2 (1 + q^2)(1 + q + q^2)}$$

A comparison with the triangular duel and the first square duel [2] for $p \to 0$, $q \to 1$ gives

$$V_B(2, 1) = \frac{1}{2} \frac{p^2 q(1 + q^2)}{(1 - q^2)(1 - q^3)} = \frac{q(1 + q^2)}{2(1 + q)(1 + q + q^2)} \xrightarrow[q \to 1]{} 1/6$$

and $V_B(2, 2) \xrightarrow[q \to 1]{} 1/2$ which are the same limits as the one obtained from Equation 29 and 37 in [2].

Table 1 gives some results for $V_B(M, N, p_A, p_B)$.

TABLE 1 — (x 10^4)

M	N	p_A / p_B	0.3 / 0.3	0.3 / 0.5	0.5 / 0.3	0.5 / 0.5	0.7 / 0.5	0.5 / 0.7	0.7 / 0.7
1	1		5000	6538	3462	5000	3824	6176	5000
2	2		5166	7307	3100	5317	3850	6873	5447
3	3		5678	8227	3405	6418	5081	8343	7386
3	5		9634	9982	8869	9913	9805	9998	9994
5	3		1292	3806	0368	1780	0997	3907	2832
5	5		7258	9614	5118	8940	8359	9920	9848
5	7		9831	9999	9422	9994	9986	10000	10000
7	5		3418	7843	1626	6060	5075	9090	8629
7	7		8850	9978	7538	9919	9853	10000	10000
10	10		9900	10000	9708	10000	10000	10000	10000

It should be noted that if $p_A = p_B = p$ and $M = N$ then $V_B \geqslant \frac{1}{2}$ and increases with p or with M. We conclude: *Parallel firing is better.*

No simple relationship exists in the case $p_A \neq p_B$. Neither Mp_A vs Np_B, nor $M^2 p_A$ vs $N^2 p_B$ are sufficient to decide if $V_B > \frac{1}{2}$.

4. SIMPLE DUEL. K HITS NECESSARY TO DESTROY

Let $M = N = 1$ and let $V_{B|A}(K_A, K_B)$ denote the probability that B wins the simple duel if A starts firing and K_A hits are necessary to destroy A and K_B for B.

It is evident that

$$V_{B|A}(K_A, K_B) = p_A V_{B|B}(K_A, K_B - 1) + q_A V_{B|B}(K_A, K_B)$$

and

$$V_{B|B}(K_A, K_B) = p_B V_{B|A}(K_A - 1, K_B) + q_B V_{B|A}(K_A, K_B).$$

This gives

$$(4.1)\qquad (1-q_A q_B)\,V_{B|A}(K_A, K_B) - p_A p_B\,V_{B|A}(K_A-1, K_B-1) - p_A q_B\,V_{B|A}(K_A, K_B-1) - q_A p_B\,V_{B|A}(K_A-1, K_B) = 0.$$

In order to solve this difference equation, following Boole [5], let us define

$$x = K_A,\ y = K_B$$
$$u_{x,y} = V_{B|A}(x-1, y-1)$$
$$D_x u = u_{x+1,y} \text{ and } D_y u = u_{x,y+1}.$$

Substituting these into Equation (4.1) we get

$$[(1-q_A q_B)D_x D_y - p_A p_B - p_A q_B D_x - q_A p_B D_y]u = 0.$$

Let $D_y = a$.

$$((1-q_A q_B)a - p_A q_B)D_x u = p_B(aq_A + p_A)u$$

which gives

$$u = p_B^x(p_A + q_A D_y)^x[(1-q_A q_B)D_y - p_A q_B]^{-x}\,\emptyset(y)$$

where $\emptyset(y)$ is arbitrary. Then,

$$u = p_b^x\left(\sum_{i=0}^{x} p_A^i q_A^{x-i} D_y^{x-i}\right)(1-q_A q_B)^{-x} D_y^{-x}\left(1 - \frac{p_A q_B}{1-q_A q_B} D_y^{-1}\right)^{-x}\emptyset(y).$$

Since $D_y^{-x}\emptyset(y) = \emptyset(y-x)$ and

$$\left(1 - \frac{p_A q_B}{1-q_A q_B} D_y^{-1}\right)^{-x} = \sum_{j=0}^{\infty}\binom{x+j-1}{j}\left(\frac{p_A q_B}{1-q_A q_B}\right)^j D_y^{-j},$$

we get

$$u = \left(\frac{p_B}{1-q_A q_B}\right)^x \sum_{i=0}^{x}\sum_{j=0}^{\infty}\binom{x}{i}\binom{x+j-1}{j} p_A^{i+j} q_A^{x-i} q_B^j (1-q_A q_B)^{-j}\,\emptyset(y-i-j).$$

Taking into account that

$$V_{B|A}(1, 1) = \frac{p_B q_A}{1-q_A q_B}$$

a good choice for $\emptyset(t)$ is

$$\emptyset(t) = 1 \text{ if } t > 0$$
$$= 0 \text{ if } t \leqslant 0.$$

Defining $r = \min(K_A, K_B - 1)$ the solution becomes

$$(4.2)\qquad V_{B|A}(K_A, K_B) = \sum_{i=0}^{r}\sum_{j=0}^{K_B-i-1}\binom{K_A}{i}\binom{K_A+j-1}{j} p_A^{i+j} p_B^{K_A} q_A^{K_A-i} q_B^j (1-q_A q_B)^{K_A-j}$$

with

$$V_{B|A}(K_A, 0) = 0 \text{ and } V_{B|A}(0, K_B) = 1.$$

One can verify by substitution that this is a solution.

One can evaluate the other probabilities of winning by

$$V_{A|B}(K_A, K_B, p_A, p_B) = V_{B|A}(K_B, K_A, p_B, p_A),$$
$$V_{B|B}(K_A, K_B, p_A, p_B) = 1 - V_{A|B}(K_A, K_B, p_A, p_B),$$

and

$$V_{A|A}(K_A, K_B, p_A, p_B) = 1 - V_{B|A}(K_A, K_B, p_A, p_B).$$

Table 2 gives some results for $V_{B|A}(K_A, K_B, p_A, p_B)$ and $V_{B|B}(K_A, K_B, p_A, p_B)$.

TABLE 2 — ($\times 10^4$)

K_A	K_B	$p_A = .3$, $p_B = .5$		$p_A = p_B = .5$		$p_A = .5$, $p_B = .7$	
		$V_{B\|A}$	$V_{B\|B}$	$V_{B\|A}$	$V_{B\|B}$	$V_{B\|A}$	$V_{B\|B}$
1	1	5385	7692	3333	6667	4118	8235
5	3	4257	5010	1139	1728	2576	3579
5	5	8201	8630	4512	5488	7414	8381
7	5	5955	6541	1674	2266	4159	5278
7	7	8695	8981	4599	5401	7981	8669
10	10	9160	9330	4671	5329	8545	9002

This table indicates that $V_B = 1/2$ if $K_A = K_B$ and $p_A = p_B = 1/2$, V_B increases towards 1 if $K_A = K_B$ and $p_B > p_A$ and $|V_{B|A} - V_{B|B}|$ decreases if K_A and K_B increase.

An interesting comparison is to be made with the results given by Bhashyam [4]. Under an assumption of an exponential distribution for interfiring times he finds that the probability of B winning is, using our notation,

$$P(B) = 1 - I_{\frac{p_A}{p_A+p_B}}(K_B, K_A)$$

where I_x is the incomplete Beta function. The correspondance in the notations being λp for p_A, $\lambda^* p^*$ for p_B, R for K_B and R^* for K_A.

Table 3 shows at what rate a model with alternate firing converges towards Bhashyam's model.

Alternate firing gives a good approximation if p is small. In fact, consider K_A and K_B fixed and $p_A = c\,p_B$ with $p_B \to 0$.

One can show that

$$\lim V_{B|A} = \frac{1}{(1+c)^{K_A}} \sum_{j=0}^{K_B-1} \binom{K_A + j - 1}{j} \left(\frac{c}{1+c}\right)^j = \lim V_{B|B}$$

and this limit from a well known theorem is

$$1 - I_{\frac{c}{1+c}}(K_B, K_A).$$

TABLE 3 — *Rate of Convergence of V_B to $P(B)$*

P_A	p_B	K_A	K_B	V_B	$P(B)$
0.4	0.2	5	5	0.1054	
0.2	0.1			0.1265	
0.02	0.01			0.1431	
0.002	0.001			0.1447	
					0.1449
0.1	0.2	5	2	0.3391	
0.01	0.02			0.3501	
0.001	0.002			0.3511	
					0.3512
0.1	0.2	10	10	0.9491	
0.01	0.02			0.9366	
					0.9352

5. SQUARE DUEL. 2 HITS NECESSARY TO DESTROY

Let $M = N = 2$ and $K_A = K_B = 2$. One can represent the state of the two forces by (i_1, i_2, j_1, j_2) with $i_1, i_2, j_1, j_2 = 0, 1, 2$, representing the number of hits necessary to destroy. For example, (1, 1; 0, 2) means that A has 2 units that can be destroyed by one hit each and B has one unit that has been destroyed by 2 hits and one unit untouched.

All attempts to arrive at one or two difference equations have been in vain. Two equivalent approaches have been used. In the first, taking $p_A = p_B = 1/2$, and defining A, as the matrix of the transitional probabilities corresponding to the case when A fires first, and B the corresponding matrix when B fires first one obtains:

V_A by summing all the probabilities for the events $(i, j; 0, 0)$ in $\lim_{n->\infty} (AB)^n$ and V_B by summing all the probabilities for the events $(0, 0; i, j)$ in $\lim_{n->\infty} (BA)^n$.

The matrices are 29 × 29. The possible states of A are such that $i_1 \geqslant i_2$. The possible states of B are such that $j_1 \leqslant j_2$ and exclude $j_1 = j_2 = 1$ since A concentrates its fire until destruction is achieved.

Assuming the ordering $i_1 \geqslant i_2$, two variations are possible. In Case 1, when the state is $(2,1; 0, j)$ with $j = 1$ or 2 and B fires, B chooses at random among the two units of A. In Case 2, B fires on the second unit of A, which can be destroyed by one shot. We find

$$\text{In Case 1} \quad V_A = 0.5586$$
$$\text{and in Case 2} \quad V_A = 0.5396$$

In both cases *concentrated firing is better.*

The other approach consists in writing down all the equations that define the battle. For example,

$$V_{A|A}(2,2;1,2) = (1 - q_A^2)\, V_{A|B}(2,2;0,2) + q_A^2\, V_{A|B}(2,2;1,2).$$

The difference between Case 1 and 2 is seen by considering

$$V_{A|B}(2,1;0,1) = 0.5\, p_B\, V_{A|A}(1,1;0,1) + 0.5\, p_B\, V_{A|A}(2,0;0,1) + q_B\, V_{A|A}(2,1;0,1)$$

or

$$V_{A|B}(2, 1;0, 1) = p_B\, V_{A|A}(2, 0;0, 1) + q_B\, V_{A|A}(2, 1;0, 1).$$

A third variation is possible in which no ordering is assumed for the i_Ks. Only the states with $i_1 = 0$ are eliminated. In this case, B fires always upon the last unit of A but 2 states are considered

$$V_{A|B}(2, 1;0, 1) = p_B\, V_{A|A}(2, 0;0, 1) + q_B\, V_{A|A}(2, 1, 0, 1)$$

and

$$V_{A|B}(1, 2;0, 1) = p_B\, V_{A|A}(1, 1;0, 1) + q_B\, V_{A|A}(2, 1;0, 1).$$

In this case $V_A = 0.5553$ for $p_A = p_B = 0.5$.

The total system consists of 35 pairs of equations and is solved by iterations.

Table 4 gives some results for the square duel in this last case. As in the two preceeding cases, concentrated firing is better.

An extension of this last case is considered in the next section.

6. MULTIPLE FAIR DUELS. K HITS NECESSARY TO DESTROY

Let us restrict ourselves to the case of a fair duel, i.e., one such that $M = N = n$, $p_A = p_B = p$ and $K_A = K_B = K$.

All nondestroyed units of A concentrate their fire on a single unit of B, volley after volley until destruction is achieved. For the next volley they concentrate their fire on the next undestroyed unit of B.

There are $nK + 1$ possible states for B

$$\begin{array}{rl} K, K, & \ldots, K \\ K-1\ K, & \ldots, K \\ & \vdots \\ 1, K, & \ldots, K \\ 0, K, & \ldots, K \\ & \vdots \\ 0, 0, & \ldots, 0. \end{array}$$

On the other hand B spreads its fire over all units of A and all states are possible, eliminating only the destroyed units.

Since there are K^{n-j} different states with j zeros the number of possible states for A is

$$(K^{n+1} - 1)/(K - 1).$$

This means that in order to find $V_A(K, \ldots, K; K, \ldots, K)$ we will have to solve a linear system consisting of $(nK + 1)\ (K^{n+1} - 1)/(K - 1)$ pairs of equations of the form

$$V_{A|A}\,(\text{state}) = \text{linear combination of } V_{A|B}\,(\text{outcome of } A \text{ firing})$$

$$V_{A|B}\,(\text{state}) = \text{linear combination of } V_{A|A}\,(\text{outcome of } B \text{ firing}).$$

TABLE 4 — (x 10^4) *Square Duel*

$P_A = P_B = p$	$V_{A\|A}$	$V_{A\|B}$	V_A	Number of Iterations
0.999	9980	40	5010	3
0.99	9809	382	5096	3
0.95	9193	1599	5396	4
0.9	8666	2573	5620	5
0.7	7573	3850	5712	8
0.5	6781	4324	5553	13
0.3	6117	4786	5452	25
0.1	5598	5198	5398	80
0.05	5487	5292	5389	157
0.025	5434	5337	5385	303
0.02	5423	5346	5385	373
0.01	5402	5364	5383	844

Unfortunately, the number of possible states increases very rapidly. A few values are given:

		Number of States
$K = 2$	$n = 2$	35
	3	105
	4	279
	5	693
	6	1651
$K = 3$	$n = 2$	91
	3	400
	4	1573.

This, however, is much better than $(K + 1)^{2n}$, which is the number of possible states without any restrictions.

Since writing down the necessary equations is an impossible task, a computer program was written to build the equations and solve them by iteration. The main steps are:

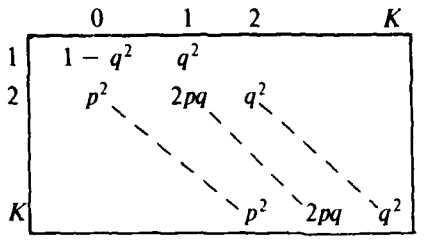
(1) define the necessary states,

(2) define $V_{A|A} = 0$ for all states
$V_{A|B} = 0$ for all states if B is not destroyed
$V_{A|B} = 1$ if B is destroyed.

These will be the initial conditions.

(3) For each state determine the number of effective units M_A and N_B. If A fires, the number of targets is $T = 1$ and the degree of concentration is $c = M_A$. If B fires, the number of targets is $T = \min(M_A, N_B)$. If $M_A \geqslant N_B$, the degree of concentration is $c = 1$ and if $N_B > M_A$, then $N_B = a\ M_A + b$ and $c_1 = a$ for $T_1 = M_A - b$ units and $c_2 = a + 1$ for $T_2 = b$ units.

(4) Let $Q_c(i, j)$ denote the probability for a unit to go from state $K = i$ to state $K = j$ if submitted to fire of concentration c. Then the matrix Q_2, for example, has the form



In general, for $i = 1, K$ and $j = 0, 1, \ldots, K$

$$Q_c(i, j) = \begin{cases} \binom{c}{i-j} p^{i-j} q^{c-i+j} & \text{for } j \neq 0 \\ 1 - \sum_{j=1}^{K} Q_c(i, j) & \text{for } j = 0. \end{cases}$$

All required matrices are constructed.

5) For each state the equation giving $V_{A|A}$ is constructed.

Let i denote the state of the target unit.

Let j denote the states of this unit after A has fired, the rest of B being unaffected.

Then,

$$V_{A|A}(A;i) = \sum_{j} Q_{M_A}(i, j)\ V_{A|B}(A; j)$$

the corresponding equation for $V_{A|B}$ is of the general form

$$V_{A|B}(i_1, i_2, \ldots, i_T; B) = \sum_{j_{e_s}} \left[\prod_{e=1}^{T} Q_{c_e}(i_e, j_{e_s}) \right] V_{A|A}(j_{1_s}, \ldots, j_{T_s}; B).$$

For example,

$$V_{A|B}(1, 2, 0, 0, 0; 1, 2, 2, 2, 2) = \sum_{\substack{j_1=0,1 \\ j_2=0,1,2}} Q_2(1, j_1)\, Q_3(2, j_2)\, V_{A|A}(j_1, j_2, 0, 0, 0; 1, 2, 2, 2, 2).$$

6) When all possible states are gone through, the last calculated value is

$$V_{A|B}(K, K, \ldots, K; K, \ldots, K).$$

It is compared, usually within 10^{-6}, to the previously calculated value and the process is iterated until convergence is achieved.

Table 5 gives results for several values of M and K. The dimension of the linear system is twice the number of states. The probability of a hit is taken as $p = 0.5$. Time is given for some cases. The computer used was a CDC6400.

The value $p = 0.5$ was chosen because time increases very fast if p decreases, as is seen from Table 4.

TABLE 5 — ($V_A \times 10^4$), *Multiple Fair duel. Exact Results*

$M = N$	K	Number of Equations	Number of Iterations	V_A	Time (in seconds)
2	2	70	13	5553	
	3	182	15	5988	
	4	378	17	6364	
	5	682	19	6661	
3	2	210	13	5537	
	3	800	15	6211	
	4	2210	17	6822	
	5	4992	19	7289	1700
4	2	558	13	5152	
	3	3146	15	6132	
	4	11594	17	6872	6141
5	2	1386	12	4429	
	3	11648	15	5931	8960

Since exact calculations of V_A become too time consuming, some results were obtained by simulation. Table 6 gives some results. The number of trials was 2000 for $p \neq 0.5$ and 6000 for $p = 0.5$, A started the duel in half the cases.

TABLE 6 — ($V_A \times 10^3$),
Multiple Fair Duel. Simulation Results

p		0.1	0.3	0.5	0.7	0.9
M	K					
2	2	550	542	561	569	564
4		574	562	522	485	403
6		583	490	351	148	4
8		580	382	120	2	0
10		562	249	11	0	0
2	3	600	600	611	622	575
4		652	632	599	628	642
6		672	606	564	494	270
8		705	554	444	181	2
10		725	482	233	6	0
2	4	586	616	642	691	778
4		715	694	685	666	722
6		774	700	653	672	651
8		796	684	608	548	205
10		797	643	524	204	1
2	5	630	646	674	708	705
4		754	753	760	748	556
6		812	786	725	665	852
8		838	777	668	698	601
10		878	740	639	594	134

We note that for large values of p the behaviour of V_A is erratic. This is due to the deterministic issue of a battle for $p = 1$ as a consequence of alternative firing. For example, if $M = 6$, $k = 2$ and A starts firing, the sequence of states is B: 022222, A: 211111, B: 002222, A: 210000, B: 000222, B wins.

Two independent estimates of the error can be made; one by comparing the results of the simulation with the calculated values in Table 5 for $p = 0.5$, $M = N = 2$ or 4 and $K = 2, 3, 4$, giving $s = 0.0093$, the other estimate is given by assuming a binomial distribution with 6000 trials giving $s = 0.0065$. To be on the safe side one can conclude that concentrated firing is better if the simulation gives $V_A \geqslant 0.519$ and parallel firing is better if the simulation gives $V_A \leqslant 0.481$. This does not take into account the bias introduced by alternate firing for "large" values of p. Since the sign of the bias is evident, one can adjust one's conclusions, for example for $M = 10$, $K = 4$ and $p = 0.5$ the observed value 0.524 is pulled down and almost certainly A wins more often than B. On the other hand for $M = 8$, $K = 3$ and $p = 0.5$ the value 0.444 is certainly pulled down and one can hardly conclude that B wins more often.

Table 7 summarizes all the results obtained.

TABLE 7 — *Better Strategy of Firing*

	Concentrated		Parallel	Border cases
$p = 0.1$	$K \geqslant 2$		$K = 1$	
$p = 0.3$	$K = 3$	$2 \leqslant M \leqslant 4$	$K = 2$ $M \geqslant 7$	$K = 2$ $M = 5$ or 6
	$K = 3$	$2 \leqslant M \leqslant 8$		$K = 3$ $M = 9$ or 10
	$K = 4, 5$	$2 \leqslant M \leqslant$ at least 10		
$p = 0.5$	$K = 2$	$2 \leqslant M \leqslant 3$	$K = 2$ $M \geqslant 5$	$K = 2$ $M = 4$
	$K = 3$	$2 \leqslant M \leqslant 6$		$K = 3$ $M \geqslant 7$
	$K = 4$	$2 \leqslant M \leqslant 10$		
	$K = 5$	$2 \leqslant M \leqslant$ at least 10		

One can conclude that concentrated firing is better if the combination of group size and hit probability does not produce a high degree of overkilling. For $K \geqslant 2$ a rough rule could be concentrate firing if $pM \leqslant K$ (the exception is $p = 0.5$, $K = 4$ and $M = 9$ or 10).

Up to this time we have compared two strategies: parallel firing and concentrated firing. In the next section we will attempt to define the concept of partial concentration.

7. MULTIPLE FAIR DUELS. PARTIAL CONCENTRATION OF FIRE

Let $M = N = n$, $p_A = p_B = p$ and $K_A = K_B = K$. Let c_X be the maximal number of non-destroyed units of A that are allowed to concentrate their fire on a single unit of B, volley after volley until destruction is achieved.

If $c_X = 1$, A uses parallel firing in the same manner as B. If $c_X = n$, A uses concentrated firing.

Under partial concentration the number of targets for A is given by the integer function

$$T_A = \left[\frac{n + c_X - 1}{c_X} \right]$$

and the number of possible states for B is

$$(n - T_A + 1)\,K^{T_A} + \frac{K^{T_A} - 1}{K - 1}$$

which reduces to $nK + 1$ for $T_A = 1$ and $(K^{n+1} - 1)/(K - 1)$ for $T_A = n$. The number of linear equations to be solved becomes

$$2\frac{K^{n+1} - 1}{K - 1}\left[(n - T_A + 1)\,K^{T_A} + \frac{K^{T_A} - 1}{K - 1}\right].$$

The value $c_X = 1$ $(T_A = n)$ was used to determine the precision of the obtained results, since $V_A = 0.5$. For $p = 0.5$ the maximum error found was 3×10^{-5} and for $p = 0.1$ it was 1×10^{-4}.

Table 8 gives some calculated results for $p = 0.5$.

TABLE 8 — $(V_A \times 10^4)$, *Partial Concentration*

$M = N$	K	c_X	Number of Equations	Number of Iterations	V_A
3	2	2	330	13	5404
	3	2	1760	15	5950
	4	2	6290	16	6412
4	2	2	930	12	5639
		3	930	13	5192
	3	2	7502	14	6304
		3	7502	15	6127
5	2	2	3906	12	5541
		3	2394	12	5311
		4	2394	12	4523

Comparing the results of Table 5 and Table 8, one sees that partial concentration with $c_X = 2$ is better than total concentration for the cases $M = 4$ and $k = 2$ or 3 and any partial concentration is better for the case $M = 5$ and $k = 2$. Further investigations are needed.

8. SUMMARY

The proposed model is an idealization of combat between small groups of individual identical firers and is very far from the very complicated process of real combat. However, it has provided, through the use of alternate firing as an expression for the interaction between opponents, some important insights into combat dynamics that could be further investigated with, for example, a high-resolution Monte Carlo simulation. It has been shown that alternate firing gives the same results for small hit probabilities as some previously developed models. It has also been shown that the relationship between the size, the hitting capacity and the resistance of the opponents is a complex one and that concentrated firing is better than alternate firing if the amount of over-killing is not too high. Moreover, some evidence suggests that partial concentration can be even more effective.

ACKNOWLEDGMENTS

This research was supported by an FIR grant from Université du Québec à Montréal. The author wishes to thank the Service de l'Informatique for its help in providing computing facilities and also the referee and an Associate Editor for their many helpful comments.

REFERENCES

[1] Ancker, C.J., Jr., "The Status of Developments in the Theory of Stochastic Duels-II," Operations Research *15*, 388-406 (1967).

[2] Ancker, C.J., Jr. and T. Williams, "Some Discrete Processes in the Theory of Stochastic Duels," Operations Research *13*, 202-216 (1965).

[3] Ancker, C.J., Jr., "Stochastic Duels with Bursts," Naval Research Logistics Quarterly *23*, 703-711 (1976).

[4] Bhashyam, N., "Stochastic Duels with Lethal Dose," Naval Research Logistics Quarterly *17*, 397-405 (1970).

[5] Boole, G., *A Treatise on the Calculus of Finite Differences* (MacMillan and Co., London, 1860).

[6] Brown, R.H., "Theory of Combat: The Probability of Winning," Operations Research *11*, 418-425 (1963).

[7] Dolansky, L., "Present State of the Lanchester Theory of Combat," Operations Research *12*, 344-358 (1964).

[8] Helmbold, R.L., "A Universal Attribution Model," Operations Research *14*, 624-635 (1966).

[9] Helmbold, R.L., "Solution of a General Non-Adaptive Many versus Many Duel Model," Operations Research *16*, 518-524 (1968).

[10] Lanchester, F.W., "Aircraft in Warfare: The Dawn of the Fourth Arm-No.V, The Principle of Concentration," Engineering 98, 422-423 (1914) (reprinted on pp. 2138-2148 of the *World of Mathematics*, J. Newman, Editor (Simon and Schuster, New York, 1956).

[11] Lanchester, F.W., *Aircraft in Warfare; the Dawn of the Fourth Arm*, (Constable and Co., London, 1916).

[12] Robertson, J.I., "A Method of Computing Survival Probabilities of Several Targets versus Several Weapons," Operations Research *4*, 546-557 (1956).

[13] Taylor, J., "Solving Lanchester-Type Equations for 'Modern Warfare' with Variable Coefficients," Operations Research *22*, 756-770 (1974).

[14] Taylor, J., "Optimal Commitment of Forces in Some Lanchester-Type Combat Models," Operations Research *27*, 96-114 (1979).

[15] Taylor, J. and S. Parry, "Force-Ratio Considerations for some Lanchester-Type Models of Warfare," Operations Research *23*, 522-533 (1975).

[16] Thompson, D.E., "Stochastic Duels Involving Reliability, Naval Research Logistics Quarterly *19*, 145-148 (1972).

[17] Williams, T., "Stochastic Duels-II," System Development Corporation Document, SP 1017/003/00, 31-61 (1963).

SPIKE SWAPPING IN BASIS REINVERSION*

R. V. Helgason and J. L. Kennington

Department of Operations Research
and
Engineering Management
Southern Methodist University
Dallas, Texas

ABSTRACT

During basis reinversion of either a product form or elimination form linear programming system, it may become necessary to swap spike columns to effect the reinversion and maintain the desired sparsity characteristics. This note shows that the only spikes which need be examined when an interchange is required are those not yet processed in the current external bump.

I. INTRODUCTION

An important component of a large scale linear programming system is the reinversion routine. This paper addresses an important ancillary technique for implementing a reinversion routine utilizing the pivot agenda algorithms of Hellerman and Rarick [5,6]. Production of factors during reinversion typically involves a left-to-right pivoting process. Unfortunately, during the left-to-right process, a proposed pivot element of a spike column may be zero, in which case columns are interchanged in an attempt to obtain a pivotable column while maintaining desired sparsity characteristics. In this paper we show that the only columns which need be considered for the interchange with a nonpivotable spike are other spikes lying to the right within the same external bump.

II. PRODUCT FORM OF THE INVERSE

Let B be any $m \times m$ nonsingular matrix. One of the most common factorizations for B^{-1} is the product form which corresponds to the method for solving a system of linear equations known as Gauss-Jordan reduction (see [3, 4]). This procedure is used to represent B^{-1} (or a row and column permutation of B^{-1}) as the product of matrices each of the form

$$Z = \begin{bmatrix} 1 & & \\ & z & \\ & & 1 \end{bmatrix} \leftarrow j\text{th row}$$

*This research was supported in part by the Air Force Office of Scientific Research under Contract Number AFOSR 77-3151.

where z is an m-component column vector, and j is called the pivot row. A few observations concerning Z are obvious.

PROPOSITION 1: Z is nonsingular if and only if $z_j \neq 0$.

PROPOSITION 2: Let β be any m-component vector having $\beta_j = 0$. Then $Z\beta = \beta$.

PROPOSITION 3: Let β be any m-component vector having $\beta_j \neq 0$, and let e^j denote the vector having jth component 1 and all other components zero.

Let $z_k = \left\{ \begin{array}{l} -\beta_k/\beta_j, \text{ if } k \neq j \\ 1/\beta_j, \text{ if } k = j \end{array} \right\}$. Then $Z\beta = e^j$.

Let $B(i)$ denote the ith column of the matrix B. Consider the following algorithm.

ALG 1: Product Form Factorization

0. *Initialization*

Interchange columns of B, if necessary, so that the first component of $B(1)$ is nonzero. Set $i \leftarrow 1$, $\beta \leftarrow B(1)$, and go to 3.

1. *Update Column*

Set $\beta \leftarrow E^{i-1} \ldots E^1 B(i)$.

2. *Swap Columns If Pivot Element Equals Zero*

If $\beta_i \neq 0$, go to 3; otherwise, there is some column $B(j)$ with $j > i$ such that the ith component of $\gamma = E^{i-1} \ldots E^1 B(j)$ is nonzero. Interchange $B(j)$ and $B(i)$ and set $\beta \leftarrow \gamma$.

3. *Obtain New Elementary Matrix*

Set

$$z_k \leftarrow \left\{ \begin{array}{l} 1/\beta_i, \text{ for } k = i \\ -\beta_k/\beta_i, \text{ otherwise,} \end{array} \right.$$

$$E^i \leftarrow \begin{bmatrix} I & & \\ & z & \\ & & I \end{bmatrix} \leftarrow i \text{ th row}$$

4. *Test for Termination*

If $i = m$, terminate; otherwise, $i \leftarrow i + 1$ and go to 1. At the termination of ALG 1, $E^m \ldots E^1$ is a row permutation of B^{-1}.

In the following two propositions we show that if in Step 2, $\beta_i = 0$, then the proposed interchange is always possible. Consider the following:

PROPOSITION 4: For $i \leqslant j$, $E^j \ldots E^1 B(i) = e^i$.

PROOF: By the construction of E^i and Proposition 3, $E^i \ldots E^1 B(i) = e^i$. By Proposition 2, $E^j \ldots E^{i+1} e^i = e^i$. So $E^j \ldots E^1 B(i) = e^i$. Using Proposition 4 we may now show the following:

PROPOSITION 5: For $2 \leqslant i \leqslant m$, let $\beta = E^{i-1} \ldots E^1 B(i)$. If $\beta_i = 0$, there is some $j > i$ such that $[E^{i-1} \ldots E^1 B(j)]_i \neq 0$.

PROOF: Suppose $[E^{i-1} \ldots E^1 B(j)]_i = 0$ for all $j > i$. By the construction of $E^1, \ldots E^{i-1}$, in ALG 1, and Proposition 1, each factor is nonsingular. Since B is nonsingular, $E^{i-1} \ldots E^1 B$ is nonsingular. By Proposition 4, $E^{i-1} \ldots E^1 B(j) = e^j$ for $1 \leqslant j \leqslant i-1$. Hence, the ith row of $E^{i-1} \ldots E^1 B$ is all zero, a contradiction.

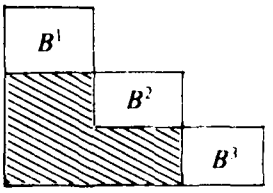
III. BUMP AND SPIKE STRUCTURE

In order to minimize the core storage required to represent the ETA file, i.e., $E^1, \ldots, E^m$, the rows and columns of B are interchanged in an attempt to place B in lower triangular form. If this can be accomplished, then the m nonidentity columns of $E^1, \ldots, E^m$, have the same sparsity structure as B. Consider the following proposition:

PROPOSITION 6: If the first $j-1$ components of $B(j)$ are zero for $j > 2$, then $E^{j-1} \ldots E^1 B(j) = B(j)$.

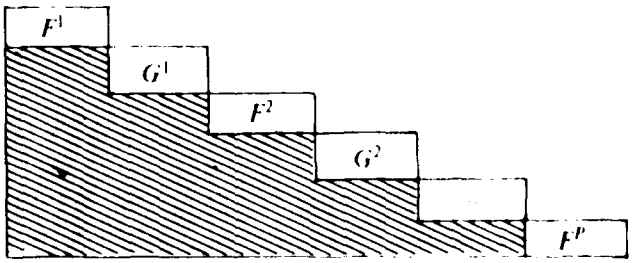
PROOF: This follows directly from successive application of Proposition 2. Therefore, if B is lower triangular, the factored representation of B^{-1} may be stored in approximately the same amount of core storage as B itself. In practice it is unneccessary to calculate the elements $1/\beta_k$ and $-\beta_i/\beta_k$ in Step 3 of ALG 1. It suffices to store k and the elements of β_i. It may prove advantageous to store $1/\beta_k$, in addition. If Proposition 6 applies for $B(k)$, then $\beta = B(k)$ and the only additional storage required is for the index k (and possibly $1/\beta_k$). Clearly, this results in substantial core storage savings compared to storing B^{-1} explicitly.

If B cannot be placed in lower triangular form, then it is placed in the form:



where B^1 and B^3 are lower triangular matrices with nonzeroes on their diagonals. We assume that if B^2 is nonvacuous, every row and column has at least two nonzero entries, so that no rearrangement of B^2 can expand the size of B^1 or B^3. B^2 is called the *bump section*, the *merit section* or the *heart section*. We further require the heart section to assume the following form:

$B^2 =$



where G^k's are either vacuous or lower triangular with nonzeroes on the diagonal. The only partitions in B having columns with nonzeroes above the diagonal are the F^k's which are called *external bumps*. The columns extending above the diagonal are called *spikes* or *spike columns*. An external bump is characterized as follows:

(i) the last column of an external bump will be a spike with a nonzero lying in the topmost row of the external bump, and

(ii) the nonspike columns have nonzero diagonal elements.

The algorithms of Hellerman and Rarick [5,6] produce such a structure for any nonsingular matrix, and we shall call a matrix having this structure an HR matrix. It should be noted that if one applies ALG 1 to an HR matrix, then the only columns which may require an interchange are spike columns. *We now prove that the only columns which need be considered for this interchange are other spikes in the same external bump.*

Consider the following result:

PROPOSITION 7: Let $B(i)$ with $i \geqslant 2$ correspond to the first column of some external bump, F^k, and let $B(j)$ be a spike in F^k. Then $E^{i-1} \ldots E^1 B(j) = B(j)$.

PROOF: Note that the first $i-1$ components of $B(j)$ are zero. Therefore, by successive application of Proposition 2, the result is proved.

Note that Proposition 6 allows one to eliminate all of the calculation required in Step 1 of ALG 1 for nonspike columns and Proposition 7 allows one to eliminate some of this calculation for spikes. We now address the issue of spike swapping. Consider the following propositions:

PROPOSITION 8: Any spike $B(j)$ which is not pivotable cannot be interchanged with a spike $B(k)$, $k > j$, from another external bump, to yield a pivotable column.

PROOF: Since $B(k)$ is from an external bump lying to the right of the external bump containing $B(j)$, $B_j(k) = 0$. By repeated application of Proposition 2, $E^{j-1} \ldots E^1 B(k) = B(k)$. Thus $B(j)$ cannot be interchanged with $B(k)$ to yield a pivotable column.

PROPOSITION 9: Any spike $B(j)$ which is not pivotable cannot be interchanged with a nonspike column $B(k)$, $k > j$, to yield a pivotable column.

PROOF: Let $B(k)$, with $k > j$ correspond to any nonspike column. From Proposition 6, $E^{j-1} \ldots E^1 B(k) = B(k)$. Since the jth component of $B(k)$ is zero, $B(j)$ cannot be interchanged with $B(k)$, to yield a pivotable column. We now present the main result of this note.

PROPOSITION 10: Any spike column $B(j)$, which is not pivotable can be interchanged with a spike, $B(k)$, with $k > j$ within the same external bump, to yield a pivotable column.

PROOF: If $B(j)$ is not pivotable, then by Proposition 5 there exists a column $B(k)$ with $k > j$ which is pivotable. By Proposition 8, $B(k)$ cannot be a spike from a different external bump. By Proposition 9, $B(k)$ cannot be a nonspike. Hence $B(k)$ must be a spike from the same external bump.

In practice, the zero check in step 2 is replaced by a tolerance check. Discussions of practical tolerance checks may be found in Benichou [1], Clasen [2], Orchard-Hays [7], Saunders [8], Tomlin [9], and Wolfe [10].

REFERENCES

[1] Benichou, M., J. Gauther, G. Hentges, and G. Ribiere, "The Efficient Solution of Large-Scale Linear Programming Problems—Some Algorithmic Techniques and Computational Results," Mathematical Programming, *13*, 280-322 (1977).

[2] Clasen, R.J., "Techniques for Automatic Tolerance Control in Linear Programming," Communications of the Association for Computing Machinery, 9, 802-803 (1966).

[3] Forsythe, G.E. and C.B. Moler, *Computer Solution of Linear Algebraic Systems*, (Prentice-Hall, Englewood Cliffs, New Jersey, 1967).

[4] Hadley, G. *Linear Algebra* (Addison Wesley Publishing Co., Inc., Reading, Massachusetts, 1964).

[5] Hellerman, E. and D. Rarick, "The Partitioned Preassigned Pivot Procedure (P^4)," *Sparse Matrices and Their Applications*, D. Rose and R. Willoughby, Editors, (Plenum Press, New York, New York, 1972).

[6] Hellerman, E. and D. Rarick, "Reinversion with the Preassigned Pivot Procedure," Mathematical Programming, *1*, 195-216 (1971).

[7] Orchard-Hays, W., *Advanced Linear Programming Computing Techniques*, (McGraw-Hill, New York, New York, 1968).

[8] Saunders, M.A., "A Fast, Stable Implementation of the Simplex Method Using Bartels-Golub Updating," *Sparse Matrix Computations*, 213-226, J.R. Bunch and D.J. Rose, Editors (Academic Press, New York, New York, 1976).

[9] Tomlin, J.A., "An Accuracy Test for Updating Triangular Factors," *Mathematical Programming Study 4*, M.L. Balinski and E. Hellerman, Editors, (North-Holland, Amsterdam, 1975).

[10] Wolfe, P., "Error in the Solution of Linear Programming Problems," *Error in Digital Computation*, *2*, L.B. Rall, Editor (John Wiley and Sons, Inc., New York, New York 1965).

AN ALTERNATIVE PROOF OF THE IFRA PROPERTY OF SOME SHOCK MODELS*

C. Derman and D. R. Smith

Columbia University
New York, New York

ABSTRACT

Let $\bar{H}(t) = \sum_{k=0}^{\infty} \frac{e^{-A(t)}A(t)^k}{k!} \bar{P}(k)$, $0 \leqslant t < \infty$, where $A(t)/t$ is nondecreasing in t, $\{\bar{P}(k)^{1/k}\}$ is nonincreasing. It is known that $H(t) = 1 - \bar{H}(t)$ is an increasing failure rate on the average (IFRA) distribution. A proof based on the IFRA closure theorem is given. $H(t)$ is the distribution of life for systems undergoing shocks occurring according to a Poisson process where $\bar{P}(k)$ is the probability that the system survives k shocks. The proof given herein shows there is an underlying connection between such models and monotone systems of independent components that explains the IFRA life distribution occurring in both models.

1. INTRODUCTION

In Barlow and Proschan [1, p. 93] a fairly general damage model is considered. A device is subject to shocks occurring in time according to a Poisson process with rate λ. The damage caused by shocks is characterized by a sequence of numbers $\{\bar{P}(k)\}$, where $\bar{P}(k)$ is the probability that the device will survive k shocks. The $\bar{P}(k)$'s as shown in [1] can arise in different models. For example, the damage caused by the ith shock can be assumed to be a nonnegative random variable X_i, where $X_1, X_2 \ldots$ are independent and identically distributed; failure of the device occurs at the kth shock if $\sum_{i=1}^{k} X_i$, the cumulative damage, exceeds a certain threshold. In this case $\bar{P}(k) = Pr\left\{\sum_{i=1}^{k} X_i \leqslant y\right\}$, where y is the threshold. Ross [2] has failure occurring when some nondecreasing symmetric function $D(X_1, \ldots, X_n)$ first exceeds a given threshold; i.e., $D(X_1, \ldots, X_k)$ is a generalization of $\sum_{i=1}^{k} X_i$. Here, $\bar{P}(k) = Pr\{D(X_1, \ldots X_k) \leqslant y\}$.

Let $\bar{H}(t)$ denote the probability that the device survives in the interval $[0, t]$. Then

$$\bar{H}(t) = \sum_{k=0}^{\infty} \frac{e^{-\lambda t}(\lambda t)^k}{k!} \bar{P}(k), \quad 0 \leqslant t < \infty.$$

In Barlow and Proschan [1] (Theorem 3.6 p. 93) it is proven that if $\{\bar{P}(k)^{1/k}\}$ is a nonincreasing sequence then $H(t) = 1 - \bar{H}(t)$ is always an increasing failure rate on the average (IFRA)

*Work supported in part by the Office of Naval Research under Contract N0014-75-0620 and the National Science Foundation under Grant No. MCS-7725-146 with Columbia University

distribution function; i.e., $\frac{-\log \bar{H}(t)}{t}$ is nondecreasing in t.

Ross [2], generalizes by allowing the Poisson process of successive shocks to be nonhomogeneous with rate function $\lambda(t)$ such that

$$\frac{A(t)}{t} = \frac{\int_0^t \lambda(s)ds}{t}$$

is nondecreasing in t. That is, the same assertion can be made when $\bar{H}(t)$ is given by

$$\bar{H}(t) = \sum_{k=0}^{\infty} \frac{e^{-A(t)}A(t)^k}{k!} \bar{P}(k), \quad 0 \leqslant t < \infty. \tag{1}$$

The proof given in [1] is based on total positivity arguments. Ross's technique for proving the IFRA result is obtained by making use of recent results [3] pertaining to what he calls increasing failure rate average stochastic processes.

Our proof below shows that all such results are a consequence of one of the central theorems of reliability theory, the IFRA Closure Theorem ([1] p. 83). This theorem asserts that a monotone system composed of a finite number of independent components, each of which has an IFRA life distribution, has itself an IFRA distribution.

It is remarked in [1, p. 91] that the coherent (or monotone) system model and the shock models under consideration are widely diverse models for which the IFRA class of distribution furnishes an appropriate description of life length, thus reenforcing the importance of the IFRA class to reliability theory. The implication of our proof is that the models are not as widely diverse as supposed.

The idea of the proof is the construction of a monotone system (of independent components, each of which has the same IFRA life distribution) whose life distribution *approximates* $H(t)$. The proof is completed by allowing the number of components in the system to increase in an appropriate way so that the approximating life distributions converge to $H(t)$; the IFRA property being preserved in the limit.

2. APPROXIMATING SYSTEMS APPROACH

For each m, $m = 1, 2 \ldots$ let $S_{m,n}$, $n = 1, 2, \ldots$ be a monotone system of n *independent* components. Let

$$\bar{P}_{m,n}(k) \equiv Pr\{\text{no cut set is formed} \mid \text{exactly } k \text{ components of } S_{m,n} \text{ are failed}\} \tag{1}$$

where all of the n components are equally likely to fail. (A cut set is a set of components such that if all components of the set fail, the system does not function). Assume

$$\bar{P}_{m,n}(k) = 0, \text{ if } k > m \text{ for every } n, \tag{2}$$

$$\lim_{n\to\infty} \bar{P}_{m,n}(k) = \bar{P}_m(k), \text{ for every } k \tag{3}$$

$$\lim_{m\to\infty} \bar{P}_m(k) = \bar{P}(k), \text{ for every } k. \tag{4}$$

We can state

THEOREM 1: If $A(t) \geq 0$, $\frac{A(t)}{t}$ is nondecreasing and (2), (3) and (4) hold, then $H(t) = 1 - \bar{H}(t)$ given by (1) is IFRA.

PROOF: Assume every component in $S_{m,n}$ is independent with life distribution $L(t) = 1 - e^{-A(t)/n}$. Then every component has an IFRA distribution. Let $Q_{m,n}(k,t)$ denote the probability that exactly k units fail within $[0, t]$. That is

$$Q_{m,n}(k,t) = \binom{n}{k}\left(1 - e^{\frac{-A(t)}{n}}\right)^k \left(e^{\frac{-A(t)}{n}}\right)^{n-k}. \tag{5}$$

Let $\bar{H}_{m,n}(t)$ denote the probability that $S_{m,n}$ works for at least t units of time, then

$$\bar{H}_{m,n}(t) = \sum_{k=0}^{m} Q_{m,n}(k,t)\bar{P}_{m,n}(k). \tag{6}$$

By the IFRA Closure Theorem, $\bar{H}_{m,n}(t)$ is IFRA.

However,

$$\begin{aligned} \bar{H}_m(t) &= \lim_{n\to\infty} H_{m,n}(t) \\ &= \sum_{k=0}^{m} \lim_{n\to\infty} Q_{m,n}(k,t)\ \bar{P}_m(k) \\ &= \sum_{k=0}^{m} \frac{e^{-A(t)}A(t)^k}{k!}\bar{P}_m(k) \\ &= \sum_{k=0}^{\infty} \frac{e^{-A(t)}A(t)^k}{k!}\bar{P}_m(k) \end{aligned} \tag{7}$$

by (2), the Poisson limit of binomial probabilities, and (3). Since the IFRA property is preserved in the limit, $\bar{H}_m(t)$ is IFRA. That is, since $\bar{H}_m(t) = \lim_{n\to\infty} \bar{H}_{m,n}(t)$ and $-(\log \bar{H}_{m,n}(t))/t$ is nondecreasing in t, then so is $-(\log \bar{H}_m(t))/t$. However,

$$\begin{aligned} \lim_{m\to\infty} \bar{H}_m(t) &= \sum_{k=0}^{\infty} \frac{e^{-A(t)}}{k!} A(t)^k \lim_{m\to\infty} \bar{P}_m(k) \\ &= \sum_{k=0}^{\infty} e^{-A(t)}\ \frac{A(t)^k}{k!}\bar{P}(k) \\ &= \bar{H}(t). \end{aligned}$$

Since again the IFRA property is preserved in the limit, it follows that $\bar{H}(t)$ is IFRA, proving the theorem.

We emphasize that the IFRA Closure Theorem is invoked only to show that that $H_{m,n}(t)$ is IFRA. The condition that $A(t)/t$ is nondecreasing is needed so that all components of $S_{m,n}$ have an IFRA distribution.

3. APPLICATION OF THEOREM

The condition that $\{\bar{P}(k)^{1/k}\}$ is a nonincreasing sequence is not used in the proof nor does it appear in the statement of Theorem 1. That the condition is implicit is due to a recent remarkable result of Ross, Shashahani and Weiss [4] that $\{\bar{P}(k)^{1/k}\}$ is necessarily nonincreasing.

To apply Theorem 1 for our purpose we must show

THEOREM 2: Let $\{\bar{P}(k)\}$ be any sequence such that $0 \leqslant \bar{P}(k) \leqslant 1$ and $\{\bar{P}(k)^{1/k}\}$ is nonincreasing. Then there exist the monotone systems $\{S_{m,n}\}$ such that (2), (3), and (4) hold.

PROOF: Let $\{\bar{P}(k)\}$ be any sequence with the hypothesized properties. Let F be any increasing continuous distribution function over $[0, \infty)$ and $\{y_k\}$ the nonincreasing sequence of nonnegative numbers such that

$$F(y_k) = \bar{P}(k)^{1/k}, \quad k = 1, 2, \ldots$$

For each $m (m < n)$ let $S_{m,n}$ be a set of n components, $i = 1, \ldots, n$. The cut sets are constructed in the following way. The ith component has an associated value x_i, $i = 1, \ldots, n$ where the values are assigned so that

$$\#\{i | x_i \leqslant x\} = [n\ F(x)], \quad 0 \leqslant x \leqslant y_1,$$
$$= n, \quad x > y_1,$$

where # means "number of" and $[\ ]$ is the greatest integer designator. Every set of k components is a cut set if $k > m$; if $k \leqslant m$ a set $(i_1, \ldots, i_k)$ of components is a cut set if and only if

$$\max (x_{i_1}, \ldots, x_{i_k}) > y_k.$$

Since $\{y_k\}$ is nonincreasing, $S_{m,n}$ is, indeed, a monotone set. But here,

$$\bar{P}_{m,n}(k) = \prod_{i=0}^{k-1} \frac{[nF(y_k)] - i}{n - i}, \quad k \leqslant m$$
$$= 0 \qquad\qquad , \quad k > m.$$

Thus,

$$\bar{P}_m(k) = \lim_{n \to \infty} \bar{P}_{m,n}(k)$$
$$= \begin{cases} F^k(y_k), & \text{if } k \leqslant m \\ 0 \qquad , & \text{if } k > m \end{cases}$$

and

$$\lim_{m \to \infty} \bar{P}_m(k) = F^k(y_k)$$
$$= \bar{P}(k), \quad k = 1, 2, \ldots .$$

This proves Theorem 2.

Theorems one and two yield the slightly more general version of Theorem 3.6 [1, p. 93].

The Ross [2] generalization follows by defining the cut sets to be determined by a nondecreasing symmetric function $D(x_1, \ldots, x_k)$; i.e., a set $i_1, \ldots, i_k$ of components is a cut set of $S_{m,n}$ if $k > m$ or, if $k \leqslant m$, when $D(x_{i_1}, \ldots, x_{i_k}) > y$, a given threshold value. From the construction of Theorem 2, Theorem 1 and the result referred to in [4] it follows that the sequence $\{\bar{P}(k)\}$ of this model satisfies the monotonicity condition. For the special case of $D(X_1, \ldots, X_k) = \sum_{i=1}^{k} X_i$, it is known that the sequence $\{\bar{P}(k)^{1/k}\}$ is nonincreasing (see [1] p. 96).

REFERENCES

[1] Barlow, R. and F. Proschan, *Statistical Theory of Reliability and Life Testing, Probability Models*, (Holt, Rinehart and Winston, New York, 1975).

[2] Ross, S.M., "Generalized Poisson Shock Model," Technical Report, Department of Industrial Engineering and Operations Research, University of California, Berkeley, California (1978).

[3] Ross, S.M., "Multivalued State Component Systems," Annals of Probability (to appear).

[4] Ross, S.M., M. Shashahani and G. Weiss, "On the Number of Component Failures in Systems whose Component Lives are Exchangeable," Technical Report, Department of Industrial Engineering and Operations Research, University of California, Berkeley, California.

NEWS AND MEMORANDA

Defense Systems Management College
Military Reservist Utilization Program

Military reservists from all U.S. Services now have a unique opportunity for a short tour at the Defense Systems Management College, Ft. Belvoir Virginia. By volunteering for the Reservist Utilization Program, an individual can increase proficiency training, maintain currency in DOD Research, Development & Acquisition Policy, contribute to the development and formulation of concepts that may become the bases of future DOD policy and help solve critical problems facing the acquisition community.

Once accepted for the program, a reservist may be assigned to one of three areas: research, education or operations. As a research associate, the individual researches and analyzes an area compatible with his training and experience. Many reservists in this category currently assist in the preparation of material for a comprehensive textbook on systems acquisition. The text will be used at DSMC by the faculty and students as well as by the systems acquisition community. As an academic consultant, a reservist provides special assistance to the College faculty by reviewing course material in his area of expertise and researching and developing training materials. In the operations/administration category, reservists administer the program by recruiting other reservists for the program, processing these reservists, and maintaining files and records.

Because of the complexity and broad scope of the systems acquisition business, the Reservist Utilization Program requires a large number of reservists from many diverse career fields. Some examples of career fields used include: engineering, procurement, manufacturing, legal, financial, personnel, administration and logistics. Reservists whose reserve duty assignments are not in these types of career fields, but who have civilian experience in these areas, are also urged to apply.

Many reservists perform their annual tours with the Reservist Utilization Program office. Others perform special tours of active duty or "mandays." When tour dates are determined and coordinated with your organization and the RUP office, submit the proper forms through your reserve organization at least 45 days prior to the tour date for an annual tour or 60 days for a special tour.

To apply for active duty or to get additional information, telephone Professor Fred E. Rosell, Jr. at commerical (703) 664-5783 or AUTOVON 354-5783. Reservists outside of Virginia may call on toll-free number (800) 336-3095 ext. 5783.

List of Referees

The Editors of the Naval Research Logistics Quarterly are grateful to the following individuals for assisting in the review of articles prior to publication.

S. Agnihotri
S. C. Albright
B. C. Archibald
H. Ascher
K. R. Baker
J. W. Barnes
F. M. Bass
H. W. Block
S. L. Brodsky
D. Butler
A. V. Cabot
M. D. Chipman
L. Cooper
E. Denardo
C. Derman
J. Dinkel
R. Ehrhardt
S. E. Elmaghraby
J. D. Esary
J. Falk
H. Feingold
M. J. Fischer
M. L. Fisher
J. C. Fisk
R. Fleming
J. O. Flynn
S. Gass
D. P. Gaver
M. Geisler
A. Geoffrion
N. D. Glassman
S. C. Graves
P. Gray
F. L. Gunther
D. Guthrie
C. M. Harris
T. C. Harris
A. Hax
P. Heidelberger
D. P. Heyman
A. J. Hoffman
P. Q. Hwang
E. Ignall
P. Jacobs
A. J. Kaplan
U. Karmarkar
A. R. Kaylan
J. L. Kennington
P. R. Kleindorfer
D. Klingman
J. E. Knepley
K. O. Kortanek
D. Kreps
W. K. Kruse
G. J. Lieberman
S. A. Lippman
D. Luenberger
R. L. McGill
W. H. Marlow
C. Marshall
K. T. Marshall
M. Mazumder
P. McKeown
K. Mehrotra
C. B. Millham
D. Montgomery
R. C. Morey
J. G. Morris
J. A. Muckstadt
S. Nahmias
M. F. Neuts
I. Olkin
J. Orlin
S. S. Panwalkar
J. H. Patterson
M. Posner
D. Reedy
H. R. Richardson
E. E. Rosinger
S. M. Ross
H. M. Salkin
R. L. Scheaffer
B. Schmeiser
P. K. Sen
J. Sethuraman
M. L. Shooman
M. Shubik
D. O. Siegmund
E. Silver
N. D. Singpurwalla
R. Soland
Henry Solomon
Herbert Solomon
R. M. Stark
L. D. Stone
W. Szwarc
H. A. Taha
J. G. Taylor
G. Thompson
W. E. Vesley
H. M. Wagner
A. R. Washburn
C. C. White
T. M. Whitin
J. D. Wiest
J. W. Wingate
R. T. Wong
M. H. Wright
S. Zacks

CORRIGENDUM:
STOCHASTIC CONTROL OF QUEUEING SYSTEMS

Dr. A. Laurinavicius of the Institute of Physical and Technical Problems of Energetics, Academy of Sciences, Lithuania, USSR, has pointed out an error in the statement of Theorem 1 of this paper [1]. The expression for the generator given there is valid only for $x > 0$, and a different expression holds for $x = 0$, the proof for this case being similar. Moreover, the domain of the generator can be extended. The correct statement is as follows.

THEOREM 1: Let the function $f(t,x)$ be continuous and such that the directional derivatives

$$D_{\vec{P}} f(t,x) = \lim_{h \to 0+} \frac{f(t+h, x-h) - f(t,x)}{h} \quad (x > 0) \tag{1}$$

$$D_{\vec{Q}} f(t,0) = \lim_{h \to 0+} \frac{f(t+h,0) - f(t,0)}{h} = \frac{\partial +}{\partial t} f(t,0) \tag{2}$$

where $\vec{P} = (1, -1)$ and $\vec{Q} = (1,0)$, exist, be continuous from one side and bounded. Then the infinitesimal generator of the semigroup $\{T_t\}$ is given by

$$\begin{aligned} Af(t,x) &= D_{\vec{P}} f(t,x) - \lambda f(t,x) + \lambda \int_0^\infty f(t, x+v) B(dv) \quad \text{for } x > 0 \\ &= D_{\vec{Q}} f(t,0) - \lambda f(t,0) + \lambda \int_{0-}^\infty f(t,v) B(dv) \quad \text{for } x = 0. \end{aligned} \tag{3}$$

As a consequence of this error the example of Section 3 does not lead to the stated result. A correct example is provided by the following. Let $r(t)$, the revenue per unit time, and $c(t)$, the operating cost per unit time, be given by

$$r(t) = r \text{ for } 0 \leqslant t \leqslant t_0, \text{ and } = 0 \text{ for } t > t_0$$

$$c(t) = c_1 \text{ for } 0 \leqslant t \leqslant t_0, \text{ and } = c_2 \text{ for } t > t_0.$$

The profit from operating the system up to a time T is given by $f(T, W_T)$, where

$$f(t,x) = r \min(t, t_0) - c_1 t_0 - c_2 \max(0, t + x - t_0). \tag{4}$$

This leads to the following correct version of Theorem 3.

THEOREM 3: Let $W_0 = w < t_0$ and assume that

$$\lambda c_2 \int_{t_0 - w}^\infty [1 - B(v)] dv < r < \lambda c_2 \beta \tag{5}$$

where β is the mean service time. Then the optimal time is given by

$$T_a = \inf\{t > 0 : t + W_t \geqslant a\} \tag{6}$$

where a is the unique solution of the equation

$$\lambda c_2 \int_{t_0 - a}^\infty [1 - B(v)] dv = r. \tag{7}$$

PROOF: It is found that for $x \geqslant 0$

$$\int_{0-}^{\infty} [f(t,x+v) - f(t,x)]B(dv) = -c_2 \int_{(t_0-t-x)^+}^{\infty} [1 - B(v)]dv$$

where $(t_0 - t - x)^+ = \max(0, t_0 - t - x)$. Also,

$$D_{\vec{P}} f(t,x) = r \text{ for } t < t_0, \text{ and } = 0 \text{ for } t \geqslant t_0 \ (x > 0)$$

$$D_{\vec{Q}} f(t,0) = r \text{ for } t < t_0, \text{ and } = -c_2 \text{ for } t \geqslant t_0.$$

Therefore, the generator in this case is given by

$$\begin{aligned} Af(t,x) &= r - \lambda c_2 \int_{(t_0-t-x)^+}^{\infty} [1 - B(v)]dv \text{ for } t < t_0, \ x \geqslant 0 \\ &= -\lambda_2 \beta \quad \text{for } t \geqslant t_0, \ x > 0 \\ &= -c_2 - \lambda c_2 \beta \quad \text{for } t \geqslant t_0, \ x = 0. \end{aligned} \tag{8}$$

In applying Theorem 2 we note that $Af(t,x) < 0$ for $t \geqslant t_0$, $x \geqslant 0$, so it suffices to consider $Af(t,x)$ for $t < t_0$, $x \geqslant 0$. We can write

$$Af(t,x) = \phi(t+x) \text{ for } t < t_0, \ x \geqslant 0,$$

where

$$\phi(t) = r - \lambda c_2 \int_{(t_0-t)^+}^{\infty} [1 - B(v)]dv. \tag{9}$$

We have

$$\phi(0) = r - \lambda c_2 \int_{t_0}^{\infty} [1 - B(v)]dv > r - \lambda c_2 \int_{t_0-w}^{\infty} [1 - B(v)]dv > 0$$

$$\phi(t_0) = r - \lambda c_2 \beta < 0$$

on account of (5). Also, $\phi(t)$ is a decreasing function of t. Therefore, there exists a unique value a such that $\phi(t) > 0$ for $0 \leqslant t < a$ and $\phi(t) < 0$ for $a < t \leqslant t_0$. Since $\phi(t) \leqslant 0$ for $t \geqslant t_0$, we have $\phi(t) \leqslant 0$ for $t \geqslant a$. This means that $Af(t,x) \leqslant 0$ for $t + x \geqslant a$, so the set R of Theorem 2 is given by $R = \{(t,x): t + x \geqslant a\}$, and the time of the first visit to R is given by (6). Since the process $t + W_t$ is monotone nondecreasing with probability one, the set R is closed. Moreover, $T_a \leqslant a$ with probability one and also $E(T_a) < \infty$. Thus, the conditions of Theorem 2 are satisfied, and T_a is optimal at $W_0 = w$, as was required to be proved.

A particular case: Let $B(x) = 1 - e^{-\mu x}$ $(x \geqslant 0, \ 0 < \mu < \infty)$. The conditions (5) reduce to

$$w < t_0 - \frac{1}{\mu} \log \left(\frac{\lambda c_2}{\mu r} \right) < t_0 \tag{10}$$

and the Equation (7) gives

$$a = t_0 - \frac{1}{\mu} \log \left(\frac{\lambda c_2}{\mu r} \right). \tag{11}$$

On account of (11) we have $a > w$.

REFERENCE

[1] Prabhu, N.U., "Stochastic Control of Queueing Systems," Naval Research Logistics Quarterly *21*, 411-418 (1974).

N.U. Prabhu
Cornell University

INDEX TO VOLUME 27

INFORMATION FOR CONTRIBUTORS

The NAVAL RESEARCH LOGISTICS QUARTERLY is devoted to the dissemination of scientific information in logistics and will publish research and expository papers, including those in certain areas of mathematics, statistics, and economics, relevant to the over-all effort to improve the efficiency and effectiveness of logistics operations.

Manuscripts and other items for publication should be sent to The Managing Editor, NAVAL RESEARCH LOGISTICS QUARTERLY, Office of Naval Research, Arlington, Va. 22217. Each manuscript which is considered to be suitable material for the QUARTERLY is sent to one or more referees.

Manuscripts submitted for publication should be typewritten, double-spaced, and the author should retain a copy. Refereeing may be expedited if an extra copy of the manuscript is submitted with the original.

A short abstract (not over 400 words) should accompany each manuscript. This will appear at the head of the published paper in the QUARTERLY.

There is no authorization for compensation to authors for papers which have been accepted for publication. Authors will receive 250 reprints of their published papers.

Readers are invited to submit to the Managing Editor items of general interest in the field of logistics, for possible publication in the NEWS AND MEMORANDA or NOTES sections of the QUARTERLY.